

ARCHIVE
for
RATIONAL MECHANICS
and
ANALYSIS

Edited by
C. TRUESDELL

Volume 1, Number 5



SPRINGER-VERLAG
BERLIN-GÖTTINGEN-HEIDELBERG
(Postverlagsort Berlin • 7. 8. 1958)

OCT 14 '58
UNIVERSITY OF HAWAII
LIBRARY

Mechanicam vero duplicem Veteres constituerunt: Rationalem quae per Demonstrationes accurate procedit, & Practicam. Ad practicam spectant Artes omnes Manuales, a quibus utique Mechanica nomen mutuata est. Cum autem Artifices parum accurate operari soleant, fit ut Mechanica omnis a Geometria ita distinguatur, ut quicquid accuratum sit ad Geometriam referatur, quicquid minus accuratum ad Mechanicam. Attamen errores non sunt Artis sed Artificum. Qui minus accurate operatur, imperfectior est Mechanicus, & si quis accuratissime operari posset, hic foret Mechanicus omnium perfectissimus.

NEWTON

La généralité que j'embrasse, au lieu d'éblouir nos lumieres, nous découvrira plutôt les véritables loix de la Nature dans tout leur éclat, & on y trouvera des raisons encore plus fortes, d'en admirer la beauté & la simplicité.

EULER

Ceux qui aiment l'Analyse, verront avec plaisir la Méchanique en devenir une nouvelle branche ...

LAGRANGE

The ARCHIVE FOR RATIONAL MECHANICS AND ANALYSIS nourishes the discipline of mechanics as a deductive, mathematical science in the classical tradition and promotes pure analysis, particularly in contexts of application. Its purpose is to give rapid and full publication to researches of exceptional moment, depth, and permanence.

Each memoir must meet a standard of rigor set by the best work in its field. Contributions must consist largely in original research; on occasion, an expository paper may be invited.

English, French, German, Italian, and Latin are the languages of the Archive. Authors are urged to write clearly and well, avoiding an excessively condensed or crabbed style.

Manuscripts intended for the Archive should be submitted to an appropriate member of the Editorial Board.

The ARCHIVE FOR RATIONAL MECHANICS AND ANALYSIS appears in numbers struck off as the material reaches the press; five numbers constitute a volume. Subscriptions may be entered through any agent. The price is DM 96.— per volume.

Notice is hereby given that for all articles published exclusive rights in all languages and countries rest with Springer-Verlag. Without express permission of Springer-Verlag, no reproduction of any kind is allowed.

Authors receive gratis 75 offprints.

Note on a Class of Exact Solutions in Magneto-hydrodynamics

C. C. LIN

1. Introduction

The purpose of this note is to describe a class of exact solutions of the magneto-hydrodynamic equations which are applicable to various flow configurations. Some examples of these solutions have been worked out in some detail by PETER GOTTLIEB; his solutions and some others will be reported in future communications.

The class of solutions considered here is of the "layer-type", including many classical exact solutions of the Navier-Stokes equations as special cases; *e.g.*, the flow near a stagnation point in the two-dimensional and the axially symmetrical cases, and the flow of a viscous fluid over a rotating disc. It represents a further generalization of a class of solutions discussed by the present writer for the steady or unsteady motion of an ordinary viscous fluid.

2. Basic equations

For a uniform electrically conducting incompressible fluid, the equations of magneto-hydrodynamics are as follows*:

$$(2.1) \quad \frac{\partial v_i}{\partial t} + v_j \frac{\partial v_i}{\partial x_j} - \frac{\mu}{4\pi\varrho} H_j \frac{\partial H_i}{\partial x_j} = -\frac{\partial \Pi}{\partial x_i} + \nu \Delta v_i, \quad \frac{\partial v_k}{\partial x_k} = 0;$$

$$(2.2) \quad \frac{\partial H_i}{\partial t} + v_j \frac{\partial H_i}{\partial x_j} - H_j \frac{\partial v_i}{\partial x_j} = \eta \Delta H_i, \quad \frac{\partial H_k}{\partial x_k} = 0.$$

In these equations, x_i ($i=0, 1, 2$) are the space coordinates, t is the time, v_i are the components of velocity, H_i are the components of the magnetic field. The properties of the fluid are characterized by the density ϱ , the kinematic viscosity ν , the magnetic permeability μ and the electric conductivity σ , in terms of which we define the magnetic diffusivity

$$(2.3) \quad \eta = (4\pi\mu\sigma)^{-1}.$$

The quantity Π is essentially the total pressure and is defined

$$(2.4) \quad \varrho \Pi = p + \mu H^2/8\pi,$$

where p is the ordinary pressure of the fluid. The summation convention is used throughout this paper.

* Cf. COWLING (1957), p. 93

We seek a class of solutions in which the x_0 -axis occupies a special position. We allow the velocity field, the magnetic field, and the pressure gradient to depend on the coordinate x_0 and the time t in a fairly general manner, but only *linearly* on the two other coordinates x_α ($\alpha = 1, 2$). To be more specific, we assume

$$(2.5) \quad \begin{aligned} v_0 &= u_0(x_0, t) \\ v_\alpha &= u_\alpha(x_0, t) + u_{\alpha\beta}(x_0, t) x_\beta, \end{aligned}$$

$$(2.6) \quad \begin{aligned} H_0 &= h_0(x_0, t) \\ H_\alpha &= h_\alpha(x_0, t) + h_{\alpha\beta}(x_0, t) x_\beta, \end{aligned}$$

and

$$(2.7) \quad \begin{aligned} \frac{\partial \Pi}{\partial x_0} &= \tilde{\omega}_0(x_0, t) \\ \frac{\partial \Pi}{\partial x_\alpha} &= \tilde{\omega}_\alpha(x_0, t) + \tilde{\omega}_{\alpha\beta}(x_0, t) x_\beta. \end{aligned}$$

Since we must have

$$(2.8) \quad \frac{\partial}{\partial x_\alpha} \left(\frac{\partial \Pi}{\partial x_0} \right) = \frac{\partial}{\partial x_0} \left(\frac{\partial \Pi}{\partial x_\alpha} \right),$$

the functions $\tilde{\omega}_\alpha(x_0, t)$ and $\tilde{\omega}_{\alpha\beta}(x_0, t)$ must actually be independent of x_0 . The function Π then takes on the form

$$(2.9) \quad \begin{aligned} \Pi &= \frac{1}{2} \tilde{\omega}_{\alpha\beta}(t) x_\alpha x_\beta + \tilde{\omega}_\beta(t) x_\beta + \tilde{\omega}(x_0, t), \\ \tilde{\omega}_0 &= \frac{\partial \tilde{\omega}}{\partial x_0}. \end{aligned}$$

3. Equations in two independent variables

The particular choice of the form of the solutions is made so that the other terms in the basic equations (2.1) and (2.2) are also *linear* in the variables x_α ($\alpha = 1, 2$). Consider, for example, $v_i \frac{\partial H_i}{\partial x_j}$. By using (2.5) and (2.6), it is easily seen that

$$(3.1) \quad v_i \frac{\partial H_i}{\partial x_i} = v_0 \frac{\partial h_0}{\partial x_0}$$

and

$$(3.2) \quad v_j \frac{\partial H_\alpha}{\partial x_j} = \left(v_0 \frac{\partial h_\alpha}{\partial x_0} + v_\gamma h_{\alpha\gamma} \right) + \left(v_0 \frac{\partial h_{\alpha\beta}}{\partial x_0} + v_\gamma h_{\alpha\gamma} \right) x_\beta.$$

Thus, these basic equations can be satisfied by comparing coefficients of x_1 and x_2 and terms independent of them.

For convenience of reference, we list below the formulae useful for such reductions. Each of the vectors v_i and H_i is of the form

$$(3.3) \quad \{A_i\} = \{a_0, a_\alpha + a_{\alpha\beta} x_\beta\},$$

where $a_0, a_\alpha, a_{\alpha\beta}$ are functions of x_0 and t . The vector $\partial A_i / \partial t$ is then of the same form. In fact

$$(3.4) \quad \left\{ \frac{\partial A_i}{\partial t} \right\} = \left\{ \frac{\partial a_0}{\partial t}, \frac{\partial a_\alpha}{\partial t} + \frac{\partial a_{\alpha\beta}}{\partial t} x_\beta \right\}.$$

Also, the Laplacian of A_i is of the form

$$(3.5) \quad \{\Delta A_i\} = \left\{ \frac{\partial^2 a_0}{\partial x_0^2}, \frac{\partial^2 a_\alpha}{\partial x_0^2} + \frac{\partial^2 a_{\alpha\beta}}{\partial x_0^2} x_\beta \right\}.$$

Furthermore, if B_i is another vector of the same form as A_i , then

$$(3.6) \quad C_i = B_j \frac{\partial A_i}{\partial x_j}$$

is also of the same form; indeed,

$$(3.7) \quad \begin{aligned} c_0 &= \frac{\partial a_0}{\partial x_0}, & c_\alpha &= b_0 \frac{\partial a_\alpha}{\partial x_0} + a_{\alpha\gamma} b_\gamma, \\ c_{\alpha\beta} &= b_0 \frac{\partial a_{\alpha\beta}}{\partial x_0} + a_{\alpha\gamma} b_{\gamma\beta}. \end{aligned}$$

With the help of these formulae, it is not difficult to write down the equations for the functions of x_0 and t defined in (2.5) and (2.6).

The equation of continuity in (2.1) is satisfied if

$$(I) \quad \frac{\partial u_0}{\partial x_0} + u_{\beta\beta} = 0.$$

Similarly, the condition for the absence of free magnetic charges is satisfied if

$$(II) \quad \frac{\partial h_0}{\partial x_0} + h_{\beta\beta} = 0.$$

The equations of motion (2.1) lead to

$$(IIIa) \quad \frac{\partial u_0}{\partial t} + u_0 \frac{\partial u_0}{\partial x_0} - \nu \frac{\partial^2 u_0}{\partial x_0^2} = -\tilde{\omega}_0 + \frac{\mu}{4\pi\varrho} \left(h_0 \frac{\partial h_0}{\partial x_0} \right),$$

$$(IIIb) \quad \frac{\partial u_\alpha}{\partial t} + u_0 \frac{\partial u_\alpha}{\partial x_0} + u_{\alpha\gamma} u_\gamma - \nu \frac{\partial^2 u_\alpha}{\partial x_0^2} = -\tilde{\omega}_\alpha + \frac{\mu}{4\pi\varrho} \left(h_0 \frac{\partial h_\alpha}{\partial x_0} + h_{\alpha\gamma} h_\gamma \right),$$

$$(IIIc) \quad \frac{\partial u_{\alpha\beta}}{\partial t} + u_0 \frac{\partial u_{\alpha\beta}}{\partial x_0} + u_{\alpha\gamma} u_{\gamma\beta} - \nu \frac{\partial^2 u_{\alpha\beta}}{\partial x_0^2} = -\tilde{\omega}_{\alpha\beta} + \frac{\mu}{4\pi\varrho} \left(h_0 \frac{\partial h_{\alpha\beta}}{\partial x_0} + h_{\alpha\gamma} h_{\gamma\beta} \right).$$

Similarly, the equations (2.2) for the change of magnetic field lead to

$$(IVa) \quad \frac{\partial h_0}{\partial t} + u_0 \frac{\partial h_0}{\partial x_0} - \eta \frac{\partial^2 h_0}{\partial x_0^2} = h_0 \frac{\partial u_0}{\partial x_0},$$

$$(IVb) \quad \frac{\partial h_\alpha}{\partial t} + u_0 \frac{\partial h_\alpha}{\partial x_0} + h_{\alpha\gamma} u_\gamma - \eta \frac{\partial^2 h_\alpha}{\partial x_0^2} = h_0 \frac{\partial u_\alpha}{\partial x_0} + u_{\alpha\gamma} h_\gamma,$$

$$(IVc) \quad \frac{\partial h_{\alpha\beta}}{\partial t} + u_0 \frac{\partial h_{\alpha\beta}}{\partial x_0} + h_{\alpha\gamma} u_{\gamma\beta} - \eta \frac{\partial^2 h_{\alpha\beta}}{\partial x_0^2} = h_0 \frac{\partial u_{\alpha\beta}}{\partial x_0} + u_{\alpha\gamma} h_{\gamma\beta}.$$

4. Consistency of the system

It is necessary to examine the above system of equations to see that there is indeed the right number of equations for the unknowns specified in (2.5), (2.6) and (2.7). Clearly, there are seven unknowns in each of (2.5) and (2.6) and only one unknown, *i.e.*, $\tilde{\omega}_0(x_0, t)$, in (2.7), since $\tilde{\omega}_\alpha$ and $\tilde{\omega}_{\alpha\beta}$ are functions of t alone and are therefore expected to be specified by boundary conditions at specified values of x_0 (*e.g.* at infinity). Thus, there are *fifteen* unknowns altogether.

On the other hand, we have one equation for each of the types (I) and (II) and seven equations for each of the types (III) and (IV), making up a total of *sixteen* equations. Thus, one of them must be redundant if we are not going to have the difficulty of too many equations. Indeed, a closer examination shows that the equations (I), (II), (IVa) and (IVc) are not all independent.

If we contract (IVc), we obtain

$$\frac{\partial h_{\beta\beta}}{\partial t} + u_0 \frac{\partial h_{\beta\beta}}{\partial x_0} + h_{\beta\gamma} u_{\gamma\beta} - \eta \frac{\partial^2 h_{\beta\beta}}{\partial x_0^2} = h_0 \frac{\partial u_{\beta\beta}}{\partial x_0} + u_{\beta\gamma} h_{\gamma\beta},$$

or

$$(IVc') \quad \frac{\partial h_{\beta\beta}}{\partial t} + u_0 \frac{\partial h_{\beta\beta}}{\partial x_0} - \eta \frac{\partial^2 h_{\beta\beta}}{\partial x_0^2} = h_0 \frac{\partial u_{\beta\beta}}{\partial x_0}.$$

In this form, it is clear that, by using (I) and (II), $u_{\beta\beta}$ and $h_{\beta\beta}$ can be eliminated to give an equation in h_0 and u_0 . The resultant equation must then be consistent with (IVa). This fact can be easily verified by noting that the x_0 -derivative of (IVa), when added onto (IVc'), leads to an identity by virtue of (I) and (II).

5. Concluding Remarks

We have thus found a large class of exact solutions of the hydromagnetic equations, involving fifteen unspecified functions of the variables x_0 and t . If one restricts oneself to the steady case, the equations above reduce to ordinary differential equations which can be integrated (numerically, if necessary) under appropriate boundary conditions. Symmetry conditions also help to simplify the equations in various cases.

These simplifications are all present in the examples worked out by PETER GOTTLIEB. He considered the flow near a stagnation point both in the two-dimensional case and in the case of axial symmetry with a magnetic field parallel to a solid boundary. These examples serve to illustrate the behavior of the magnetic field as it is carried along a solid surface due to convection currents in the core of the earth. It is clear, from a consideration of the path of the fluid particles, that there can be a considerable stretching of the magnetic lines. In the ideal case of zero viscosity and zero magnetic diffusivity, the field may be expected to become infinite as the boundary is approached. It is to be expected that this tendency to increase will be counteracted by both the hydrodynamic and the magnetic diffusive effects. The resultant picture in the case of steady motion is worked out in a paper to follow. Other interesting problems may be developed within the present framework. For example, one might consider an infinite disc oscillating about its axis (taken to be the x_0 -axis) in the presence of a magnetic field in that direction. Some hydromagnetic damping is certainly to be expected. The steady rotation of a disc would be another possible problem, extending that worked out by VON KÁRMÁN & COCHRAN years ago* in the purely hydrodynamic case. Indeed, this may be combined with the problems studied by GOTTLIEB to describe the combined effect of earth rotation and convection near a stagnation point in the problem of geomagnetism. Finally, one might consider the problem of a uniform magnetic field perpendicular to the x_0 -axis

* See GOLDSTEIN (1938), p. 110.

and rotating about it (perhaps in an unsteady manner), and attempt to find out whether this would indeed provide an effective "magnetic wall" to hold an approaching fluid. It is clear from elementary considerations of Eq. (2.1) and (2.2) that a magnetic field can act in the manner of a static pressure to hold a fluid in balance, but the added effect of rotation may have the advantage of making the equilibrium more stable. These and other problems are subjects for future studies.

Acknowledgement. This work was partly sponsored by the Office of Naval Research and may be reproduced for any purpose of the United States Government.

References

- COWLING, T. G.: *Magnetohydrodynamics*. New York: Interscience 1957.
GOLDSTEIN, S.: *Modern Developments in Fluid Dynamics*. Oxford 1938.
GOTTLIEB, PETER: Some Exact Solutions in Hydromagnetics (to be published).

Massachusetts Institute of Technology,
Cambridge, Massachusetts

(Received March 31, 1958)

Work Functions in Hypo-elasticity

B. BERNSTEIN & J. L. ERICKSEN

1. Introduction

According to linear elasticity theory, the work per unit volume done by the stress in deforming a material is uniquely determined by the initial and final states of stress, regardless of what intermediate stresses it may experience. The theory of hypo-elasticity proposed by TRUESDELL [1, 2, 3] constitutes one generalization of linear elasticity theory. As will be made clear later, the work density for hypo-elastic materials generally depends on the entire stress history. In other words, it is a functional rather than a function of the stress. The main purpose of this paper is to characterize two exceptional types of hypo-elastic materials for which the work density depends only on the end states of stress. THOMAS [4, 5] has given examples of materials similar to those discussed here. The materials of type II discussed below are included in the class of materials characterized in [6]. The work of CAPRIOLI [7] suggested to us some of the results presented, though his analysis is not directly applicable to hypo-elasticity.

We begin by establishing theorems on functionals which are useful for our purpose. The remainder of the paper is devoted to the application and interpretation of these theorems in hypo-elasticity theory.

2. Functionals considered

We study functionals of the form

$$F[y^\alpha(\tau), t] = \int_0^t F_\alpha[y^\alpha(\tau)] \frac{dy^\alpha(\tau)}{d\tau} e^{\int_0^\tau G_\gamma[y^\delta(\eta)] \frac{dy^\gamma(\eta)}{d\eta} d\eta} d\tau, \quad (2.1)$$

where Greek indices run from 1 to n , the summation convention is used, F_α and G_α are specified functions of y^β , and $y^\alpha = y^\alpha(\tau)$ represents an arbitrary curve. Except when otherwise noted, it is assumed that F_α and G_α are continuously differentiable functions of y^β in some region \mathfrak{R} and that the curves considered lie in \mathfrak{R} and are continuous and piecewise continuously differentiable functions of τ in the τ -intervals considered. Results not clearly labelled as local are global, the conclusions holding throughout any subregion of \mathfrak{R} in which the hypotheses hold. For ease in writing, we abbreviate (2.1) by the expression

$$F[f^\alpha(\tau), t] = \int_0^t F_\alpha df^\alpha(\tau) e^{\int_0^\tau G_\beta df^\beta(\eta)} \quad (2.2)$$

when the curve is given by $y^\alpha = f^\alpha(\tau)$.

Lemma 1. If the curves $y^\alpha = f^\alpha(\tau)$ and $y^\alpha = g^\alpha(\tau)$ are related by

$$f^\alpha(t_1 - \tau) = g^\alpha(\tau), \quad (2.3)$$

in the τ -interval $\langle 0, t_1 \rangle$, then

$$F[f^\alpha(\tau), t_1] = -e^{-\int_0^{t_1} G_\beta df^\beta(\tau)} F[g^\alpha(\tau), t_1]. \quad (2.4)$$

Proof. In the left side of (2.4), make the changes of variable

$$\bar{\tau} = t_1 - \tau, \quad \bar{\eta} = t_1 - \eta, \quad (2.5)$$

obtaining by (2.2)

$$\begin{aligned} F[f^\alpha(\tau), t_1] &= \int_{t_1}^0 F_\alpha df^\alpha(\bar{\tau}) e^{\int_{t_1}^{\bar{\tau}} G_\beta df^\beta(\bar{\eta})} \\ &= -e^{-\int_0^{t_1} G_\beta df^\beta(\bar{\eta})} \int_0^{t_1} F_\alpha df^\alpha(\bar{\tau}) e^{\int_0^{\bar{\tau}} G_\gamma df^\gamma(\bar{\eta})} \\ &= -e^{-\int_0^{t_1} G_\beta df^\beta(\bar{\tau})} F[f^\alpha(\bar{\tau}), t_1]. \end{aligned}$$

Using (2.3) and (2.5), we then obtain (2.4).

Lemma 2. Let $y^\alpha = a^\alpha \in \mathfrak{R}$ represent any fixed point and let t_1 denote any fixed τ -value. If $F[f^\alpha(\tau), t_1] \geq 0$ for every curve $y^\alpha = f^\alpha(\tau)$ such that

$$f^\alpha(0) = f^\alpha(t_1) = a^\alpha, \quad (2.6)$$

then $F[f^\alpha(\tau), t_1] = 0$ for every such curve.

Proof. Given any curve $y^\alpha = f^\alpha(\tau)$ which satisfies (2.6), we can construct the curve

$$y^\alpha = g^\alpha(\tau) = f^\alpha(t_1 - \tau),$$

which also satisfies (2.6). By hypothesis,

$$F[g^\alpha(\tau), t_1] \geq 0, \quad F[f^\alpha(\tau), t_1] \geq 0.$$

From Lemma 1, this is impossible unless $F[f^\alpha(\tau), t_1] = 0$.

CAPRIOLI [7] established a lemma which is equivalent to our Lemma 2 when $G_\alpha \equiv 0$.

Lemma 3. Let $y^\alpha = a^\alpha$ represent any fixed point in \mathfrak{R} . A necessary and sufficient condition that, for each choice of $t^* \geq 0$,

$$F[f^\alpha(\tau), t^*] = 0 \quad \text{whenever} \quad f^\alpha(0) = f^\alpha(t^*) = a^\alpha \quad (2.7)$$

is that there exist a function $\gamma = \gamma(y^\alpha)$,

$$\gamma(a^\alpha) = 0, \quad (2.8)$$

such that, for every $t \geq 0$,

$$F[f^\alpha(\tau), t] = \gamma[f^\alpha(t)] e^{\int_0^t G_\beta df^\beta(\tau)} \quad (2.9)$$

whenever

$$f^\alpha(0) = a^\alpha. \quad (2.10)$$

Proof. From (2.8), (2.9) and (2.10), we have

$$F[f^\alpha(\tau), t^*] = \gamma[f^\alpha(t_1)] e^{\int_0^{t^*} G_\beta df^\beta(\tau)} = \gamma(a^\alpha) e^{\int_0^{t^*} G_\beta df^\beta(\tau)} = 0$$

whenever $f^\alpha(0) = f^\alpha(t^*) = 0$, so these equations imply (2.7). Suppose now that (2.7) holds. Select any point $y^\alpha = b^\alpha \in \mathfrak{R}$ and any two positive numbers t_1 and t_2 and consider any two curves

$$y^\alpha = f_N^\alpha(\tau), \quad 0 \leq \tau \leq t_N \quad (N = 1, 2)$$

such that

$$f_N^\alpha(0) = a^\alpha, \quad f_N^\alpha(t_N) = b^\alpha. \quad (2.11)$$

When $b^\alpha = a^\alpha$, we may take $t_1 = 0$. The curve $y^\alpha = f^\alpha(\tau)$, where

$$\begin{aligned} f^\alpha(\tau) &= f_1^\alpha(\tau), \quad 0 \leq \tau_1, \\ &= f_2^\alpha(t_1 + t_2 - \tau), \quad t_1 \leq \tau \leq t_1 + t_2, \end{aligned} \quad (2.12)$$

satisfies

$$f^\alpha(0) = f^\alpha(t^*) = a^\alpha, \quad t^* = t_1 + t_2.$$

From (2.7), (2.2) and (2.12),

$$\begin{aligned} 0 &= F[f^\alpha(\tau), t_1 + t_2] = \int_0^{t_1+t_2} F_\alpha df^\alpha(\tau) e^{\int_0^\tau G_\beta df^\beta(\eta)} \\ &= \int_0^{t_1} F_\alpha df_1^\alpha(\tau) e^{\int_0^\tau G_\beta df^\beta(\eta)} + \int_{t_1}^{t_1+t_2} F_\alpha df^\alpha(\tau) e^{\int_0^\tau G_\beta df^\beta(\eta)} \\ &= F[f_1^\alpha(\tau), t_1] + e^{\int_0^{t_1} G_\beta df_1^\beta(\eta)} \int_{t_1}^{t_1+t_2} F_\alpha df^\alpha(\tau) e^{\int_{t_1}^\tau G_\gamma df_2^\gamma(\eta)}. \end{aligned} \quad (2.13)$$

Introducing the changes of variable $\bar{\tau} = t_1 + t_2 - \tau$, $\bar{\eta} = t_1 + t_2 - \eta$ and using (2.12), we obtain

$$\begin{aligned} \int_{t_1}^{t_1+t_2} F_\alpha df^\alpha(\tau) e^{\int_{t_1}^\tau G_\gamma df_2^\gamma(\eta)} &= \int_{t_2}^0 F_\alpha f_2^\alpha(\bar{\tau}) e^{\int_{t_2}^{\bar{\tau}} G_\gamma df_2^\gamma(\bar{\eta})} \\ &= -e^{-\int_0^{t_2} G_\gamma df_2^\gamma(\bar{\eta})} \int_0^{t_2} F_\alpha df_2^\alpha(\bar{\tau}) e^{\int_0^{\bar{\tau}} G_\beta df_2^\beta(\bar{\eta})} \\ &= -e^{-\int_0^{t_2} G_\gamma df_2^\gamma(\eta)} F[f_2^\alpha(\tau), t_2], \end{aligned}$$

which, with (2.13), yields

$$e^{-\int_0^{t_1} G_\beta df_1^\beta(\eta)} F[f_1^\alpha(\tau), t_1] = e^{-\int_0^{t_2} G_\beta df_2^\beta(\eta)} F[f_2^\alpha(\tau), t_2]. \quad (2.14)$$

This must hold for every pair of curves satisfying (2.11), where b^α , t_1 and t_2 are assigned arbitrarily. Thus a function $\gamma(y^\alpha)$ is well defined by

$$\gamma(y^\alpha) \equiv e^{-\int_0^{t'} G_\beta df^\beta(\eta)} F[f^\alpha(\tau), t'], \quad (2.15)$$

where $f^\alpha(\tau)$ represents any curve passing through the point y^α such that $f^\alpha(0) = a^\alpha$, t' being a value of τ taken on when the curve passes through y^α . For any such curve,

$$\gamma(a^\alpha) = \gamma[f^\alpha(0)] = F[f^\alpha(\tau), 0] = 0,$$

which establishes (2.8). From (2.15), (2.9) follows immediately when (2.10) holds.

Lemma 4. *In order that there exist a function $\gamma(y^\alpha)$ and a point $y^\alpha = a^\alpha$ such that*

$$F[f^\alpha(\tau), t] = \gamma[f^\alpha(t)] e^{\int_0^t G_\beta df^\beta(\tau)} \quad (2.16)$$

whenever $f^\alpha(0) = a^\alpha$ it is necessary and sufficient that there exist a solution $\gamma(y^\alpha)$ of the system

$$\frac{\partial \gamma}{\partial y^\alpha} = -\gamma G_\alpha + F_\alpha \quad (2.17)$$

such that, at some point $y^\alpha = a^\alpha$,

$$\gamma(a^\alpha) = 0. \quad (2.18)$$

Proof. Suppose (2.16) holds. For any continuously differentiable curve $y^\alpha = f^\alpha(t)$ such that $f^\alpha(0) = a^\alpha$, $\gamma[f^\alpha(t)]$ is then a differentiable function of t and

$$\frac{d}{dt} F[f^\alpha(\tau), t] = \left[\frac{d\gamma}{dt} [f^\alpha(t)] + \gamma G_\alpha \frac{df^\alpha}{dt} \right] e^{\int_0^t G_\beta df^\beta(\tau)}. \quad (2.19)$$

From (2.2), we also have

$$\frac{d}{dt} F[f^\alpha(\tau), t] = F_\alpha \frac{df^\alpha}{dt} e^{\int_0^t G_\beta df^\beta(\tau)}, \quad (2.20)$$

so that

$$\frac{d\gamma}{dt} = -\gamma G_\alpha \frac{df^\alpha}{dt} + F_\alpha \frac{df^\alpha}{dt}, \quad (2.21)$$

Now $f^\alpha(0) = a^\alpha$ imposes no restriction on the values of $\frac{df^\alpha}{dt}$ for $t \geq 0$ or on the values of $f^\alpha(t)$ for $t > 0$. It thus follows from (2.21) that the partial derivatives $\frac{d\gamma}{dy^\alpha}$ exist and satisfy (2.17). That (2.18) must hold follows from (2.16) and the fact that $F[f^\alpha(\tau), 0] = 0$ for all curves.

Suppose conversely that (2.17) and (2.18) hold. Then, for all differentiable curves, (2.20) and (2.21) hold, from which (2.19) follows. Integrating (2.19) from $\tau = 0$ to t , then using (2.18), we obtain (2.16) whenever $f^\alpha(0) = a^\alpha$.

Lemma 5. *In order that (2.17), considered as a system of differential equations in the unknown γ , admit more than one solution, it is necessary and sufficient that there exist, at least locally, functions $A(y^\alpha)$ and $\Theta(y^\alpha)$ such that*

$$G_\alpha = \partial A / \partial y^\alpha, \quad (2.22)$$

$$F_\alpha = e^{-A} \partial \Theta / \partial y^\alpha. \quad (2.23)$$

Proof. It is easily verified that, if (2.22) and (2.23) hold, (2.17) admits the solutions $\gamma = e^{-A}(\Theta + A)$, where A is an arbitrary constant. Conversely, if (2.17)

admits two distinct solutions γ_1 and γ_2 , the difference $\gamma_1 - \gamma_2 = \Delta$ satisfies

$$\frac{\partial \Delta}{\partial y^\alpha} = -\Delta G_\alpha,$$

from which (2.22) holds with $\Delta = -\ln \Delta$. From (2.17) and (2.22),

$$F_\alpha = \frac{\partial \gamma_1}{\partial y^\alpha} + \gamma_1 \frac{\partial \Delta}{\partial y^\alpha} = e^{-\Delta} \frac{\partial}{\partial y^\alpha} \gamma_1 e^\Delta,$$

so (2.23) holds with $\Theta = \gamma_1 e^\Delta$.

3. Functionals of type I

We pay special attention to two types of functionals, the first being characterized as follows:

Definition 1. A functional of the form (2.2) is of type I provided there exists a point $y^\alpha = a^\alpha \in \mathfrak{R}$ such that $F[f^\alpha(\tau), t] \geq 0$ for any $t \geq 0$ and for every curve $y^\alpha = f^\alpha(\tau)$ such that $f^\alpha(0) = a^\alpha$.

Theorem 1. A functional is of type I if and only if there exists a point $y^\alpha = a^\alpha \in \mathfrak{R}$ and a function $\gamma(y^\alpha)$ such that

$$\gamma(y^\alpha) \geq 0, \quad \gamma(a^\alpha) = 0, \quad (3.1)$$

and such that, for any $t \geq 0$,

$$F[f^\alpha(\tau), t] = \gamma[f^\alpha(t)] e^{\int_0^t G_\beta df^\beta(\tau)} \quad (3.2)$$

whenever

$$f^\alpha(0) = a^\alpha. \quad (3.3)$$

Proof. From (3.1), (3.2) and (3.3), it follows immediately that $F[f^\alpha(\tau), t] \geq 0$ whenever (3.3) holds, so these conditions are sufficient for a functional to be of type I.

For any functional of type I, the hypotheses of Lemma 2 are satisfied for arbitrary $t_1 \geq 0$. From Lemma 2 and Lemma 3, there then exists a function such that (3.2) holds whenever (3.3) does. That γ must be non-negative follows from (3.2) and Definition 1.

Corollary 1. A functional is of type I if and only if there exists a point $y^\alpha = a^\alpha$ and a function γ such that

$$\frac{\partial \gamma}{\partial y^\alpha} = -\gamma G_\alpha + F_\alpha \quad (3.4)$$

and such that

$$\gamma(y^\alpha) \geq 0, \quad \gamma(a^\alpha) = 0. \quad (3.5)$$

Proof. This result follows immediately from Lemma 4 and Theorem 1.

Corollary 2. For any functional of type I,

$$F_\alpha(a^\beta) = 0. \quad (3.6)$$

Proof. From (3.5), γ takes on zero as a minimum value at $y^\alpha = a^\alpha$ so that $\partial \gamma / \partial y^\alpha = \gamma = 0$ when $y^\alpha = a^\alpha$. From (3.4), we then have (3.6).

Using Corollary 1, it is easy to construct functionals of type I: we can choose the functions G_α arbitrarily, select any non-negative function γ which vanishes at some point, then use (3.4) to calculate F_α .

4. Functionals of type II

The second type of functionals to be considered is described as follows:

Definition 2. A functional of the form (2.2) is a functional of type II provided that, for each $t_1 \geq 0$, $F[f^\alpha(\tau), t_1] \geq 0$ for every curve $y^\alpha = f^\alpha(\tau)$ such that $f^\alpha(0) = f^\alpha(t_1)$.

Theorem 2. If the trivial case $F_\alpha \equiv 0$ is excluded, the following three statements are equivalent:

- The functional given by (2.2) is of type II.
- For each $t_1 \geq 0$, $F[f^\alpha(\tau), t_1] = 0$ for every curve $y^\alpha = f^\alpha(\tau)$ such that $f^\alpha(0) = f^\alpha(t_1)$.
- There exist functions $\Theta(y^\alpha)$ and $\Lambda(y^\alpha)$ such that, for any $t \geq 0$ and for every curve $y^\alpha = f^\alpha(\tau)$,

$$\int_0^t G_\alpha df^\alpha(\tau) = \Lambda[f^\alpha(t)] - \Lambda[f^\alpha(0)], \quad (4.1)$$

$$F[f^\alpha(\tau), t] = \{\Theta[f^\alpha(t)] - \Theta[f^\alpha(0)]\} e^{-\Lambda[f^\alpha(0)]}. \quad (4.2)$$

Proof. That a) and b) are equivalent follows immediately from Lemma 2 and Definition 2. From (4.2), it is clear that c) implies b). It remains to show that b) implies c). When b) holds, it follows from Lemma 3 that, for each choice of a^α , there exists a function $\gamma(y^\alpha)$ such that, for any $t \geq 0$, and any curve $y^\alpha = f^\alpha(\tau)$,

$$F[f^\alpha(\tau), t] = \gamma[f^\alpha(t)] e^{\int_0^t G_\beta df^\beta(\tau)}, \quad \gamma(a^\alpha) = 0$$

whenever $f^\alpha(0) = a^\alpha$. Equivalently, there exists a function χ of pairs of values of y^α such that, for any $t \geq 0$ and any curve $y^\alpha = f^\alpha(\tau)$,

$$F[f^\alpha(\tau), t] = \chi[f^\alpha(0), f^\alpha(t)] e^{\int_0^t G_\beta df^\beta(\tau)},$$

or

$$\begin{aligned} \chi[f^\alpha(0), f^\alpha(t)] &= e^{-\int_0^t G_\beta df^\beta(\tau)} F[f^\alpha(\tau), t] \\ &= \int_0^t F_\alpha df^\alpha(\tau) e^{\int_0^\tau G_\beta df^\beta(\eta)}, \quad \chi[f^\alpha(0), f^\alpha(0)] = 0. \end{aligned} \quad (4.3)$$

Select any three points $y^\alpha = a^\alpha$, b^α and c^α in \Re and any curve $y^\alpha = f^\alpha(\tau)$ such that, for some τ values t_1 and $t_2 \geq t_1 \geq 0$, $f^\alpha(0) = a^\alpha$, $f^\alpha(t_1) = b^\alpha$ and $f^\alpha(t_2) = c^\alpha$. From (4.3), we then have

$$\begin{aligned} \chi(a^\alpha, c^\alpha) &= \int_0^{t_2} F_\alpha df^\alpha(\tau) e^{\int_0^\tau G_\beta df^\beta(\eta)} \\ &= e^{\int_0^{t_1} G_\beta df^\beta(\eta)} \int_0^{t_1} F_\alpha df^\alpha(\tau) e^{\int_0^\tau G_\gamma df^\gamma(\eta)} + \int_{t_1}^{t_2} F_\alpha df^\alpha(\tau) e^{\int_0^\tau G_\beta df^\beta(\eta)} \\ &= e^{\int_0^{t_1} G_\beta df^\beta(\eta)} F[a^\alpha, b^\alpha] + \int_0^{t_2-t_1} F_\alpha df^\alpha(\bar{\tau}) e^{\int_0^{\bar{\tau}} G_\beta df^\beta(\bar{\eta})}, \end{aligned}$$

where $\bar{\tau} = \tau - t_1$, $\bar{\eta} = \eta - t_1$. From (4.3), the last term has the value $\chi(b^\alpha, c^\gamma)$. We thus have

$$\chi(a^\alpha, c^\gamma) = e^{\int_{t_2}^{t_1} G_\beta df^\beta(\tau)} \chi(a^\alpha, b^\gamma) + \chi(b^\alpha, c^\gamma), \quad (4.4)$$

whenever

$$f^\alpha(0) = a^\alpha, \quad f^\alpha(t_1) = b^\alpha, \quad f^\alpha(t_2) = c^\alpha. \quad (4.5)$$

Since $f^\alpha(0)$ is arbitrary in (4.3),

$$\chi(y^\alpha, y^\gamma) = 0 \quad \text{for all } y^\alpha. \quad (4.6)$$

When F_α does not vanish identically, there always exist pairs of points for which $\chi \neq 0$, e.g., two distinct points on a curve satisfying $dy^\alpha/d\tau = F_\alpha$, one being a point at which $F_\alpha \neq 0$. Let a^α and b^α be two such points and consider the case where $c^\alpha = b^\alpha$ but t_1 and t_2 may differ. From (4.4), (4.5) and (4.6), we then have

$$\int_{t_2}^{t_1} G_\beta df^\beta(\tau) = 0 \quad \text{whenever } f^\alpha(t_1) = f^\alpha(t_2) = b^\alpha, \quad (4.7)$$

from which (4.4) follows immediately. From (4.1), (4.4) and (4.5), we then have, for all a^α , b^α and c^α ,

$$\chi(a^\alpha, c^\gamma) e^{A(c^\mu)} - \chi(a^\alpha, b^\gamma) e^{A(b^\mu)} = \chi(b^\alpha, c^\gamma) e^{A(c^\mu)}. \quad (4.8)$$

Setting $\star a^\alpha = 0$, $\Theta(c^\alpha) = \chi(0, c^\alpha) e^{A(c^\beta)}$, we obtain from (4.8)

$$e^{A(c^\mu)} \chi(b^\alpha, c^\gamma) = \Theta(c^\alpha) - \Theta(b^\alpha), \quad (4.9)$$

for all b^α and c^α . From (4.3), (4.1) and (4.9)

$$\begin{aligned} F[f^\alpha(\tau), t] &= e^{\int_0^t G_\alpha df^\alpha(\eta)} \chi[f^\beta(0), f^\gamma(t)] \\ &= e^{-A[f^\alpha(0)]} e^{A[f^\mu(t)]} \chi[f^\beta(0), f^\gamma(t)] \\ &= e^{-A[f^\alpha(0)]} \{\Theta[f^\beta(t)] - \Theta[f^\beta(0)]\}, \end{aligned}$$

which is (4.2). Hence b) implies c).

The functions Θ and A are not uniquely determined by (4.1) and (4.2). As is easily verified, we have

$$\begin{aligned} A(y^\alpha) - A(z^\alpha) &= A'(y^\alpha) - A'(z^\alpha), \\ e^{-A(z^\alpha)} [\Theta(y^\beta) - \Theta(z^\beta)] &= e^{-A'(z^\alpha)} [\Theta'(y^\beta) - \Theta'(z^\beta)], \end{aligned} \quad (4.10)$$

for all y^α and z^α if and only if there exist constants **A** and **B** such that

$$A = A' + \mathbf{A}, \quad \Theta = e^{\mathbf{A}} (\Theta' + \mathbf{B}). \quad (4.11)$$

Theorem 3. *The functional given by (2.2) is of type II if and only if there exist functions $A(y^\alpha)$ and $\Theta(y^\alpha)$ such that*

$$G_\alpha = \partial A / \partial y^\alpha, \quad (4.12)$$

$$F_\alpha = e^{-A} \partial \Theta / \partial y^\alpha. \quad (4.13)$$

* If $a^\alpha = 0$ is not in \mathfrak{R} , we may select any other fixed point which is.

Proof. From Theorem 2, it suffices to show that (4.1) and (4.2) hold if and only if (4.12) and (4.13) hold. Clearly (4.1) holds if and only if (4.12) holds. Differentiation of (4.2) with respect to t gives, for arbitrary continuously differentiable curves,

$$F_\alpha \frac{d f^\alpha}{dt} e^{\int_0^t G_\beta d f^\beta(\eta)} = e^{-\Lambda[f^\alpha(0)]} \frac{d}{dt} \Theta[f^\beta(t)]. \quad (4.14)$$

Using (4.1), we may write (4.14) in the form

$$F_\alpha \frac{d f^\alpha}{dt} e^{\Lambda[f^\beta(t)]} = \frac{d}{dt} \Theta[f^\beta(t)], \quad (4.15)$$

from which (4.13) follows. Conversely, if (4.13) holds, (4.15) follows for any continuously differentiable curve. Multiplying (4.15) by $e^{-\Lambda[f^\alpha(0)]}$ and using (4.1), we obtain (4.14). Integration of (4.14) yields (4.2).

From Theorem 3, Lemma 5 and Corollary 1, a functional of type II is similar to a functional of type I in that, for either type, there exists at least one function γ satisfying (3.4). For a functional of type II, it is not necessarily true that there will exist a solution of (3.4) satisfying (3.5).

Theorem 4. *For a functional to be of type I and of type II, it is necessary and sufficient that there exist functions $\Lambda(y^\alpha)$ and $\Theta(y^\alpha)$ satisfying (4.12) and (4.13) such that Θ takes on an absolute minimum at some point $y^\alpha = a^\alpha \in \mathbb{R}$.*

Proof. From Theorem 3, a functional is of type II if and only if there exist functions Λ and Θ such that (4.12) and (4.13) hold. From Corollary 1, a functional is of type I if and only if (3.4) and (3.5) hold for some point $y^\alpha = a^\alpha$. A functional is thus of both types if and only if there exist functions γ , Λ and Θ such that

$$F_\alpha = e^{-\Lambda} \frac{\partial \Theta}{\partial y^\alpha} = \frac{\partial \gamma}{\partial y^\alpha} + \gamma G_\alpha = \frac{\partial \gamma}{\partial y^\alpha} + \gamma \frac{\partial \Lambda}{\partial y^\alpha}, \quad (4.16)$$

with

$$\gamma(a^\alpha) = 0, \quad \gamma(y^\alpha) \geq 0 \quad (4.17)$$

for some point $y^\alpha = a^\alpha$. For (4.16) to hold, it is necessary and sufficient that (4.12) and (4.13) hold and that

$$\gamma = e^{-\Lambda} (\Theta + A), \quad A = \text{const.} \quad (4.18)$$

For (4.17) and (4.18) to be consistent, it is necessary and sufficient that

$$\Theta(y^\alpha) \geq \Theta(a^\alpha), \quad A = -\Theta(a^\alpha),$$

which completes the proof.

5. Hypo-elasticity equations

Following NOLL [9] and THOMAS [10, 11], we write the constitutive equations for hypo-elastic materials in the form

$$\frac{D t_{ij}}{dt} \equiv \frac{d t_{ij}}{dt} - t_{ik} w_{jk} - t_{jk} w_{ik} = A_{ijklm} d_{km}, \quad (5.1)$$

using Cartesian tensor notation, where $t_{ij} = t_{ji}$ represents the stress tensor, t the time, d/dt the material derivative, w_{jk} and d_{jk} are given in terms of the velocity

components v_i by $2w_{jk} = v_{j,k} - v_{k,j}$, $2d_{jk} = v_{j,k} + v_{k,j}$ and A_{ijkm} represents a tensor invariant or isotropic function of t_{pq} with the symmetry $A_{ijkm} = A_{jikm} = A_{jimk}$. Dt_{ij}/dt transforms as a second order tensor under arbitrary time-dependent orthogonal coordinate transformations and reduces to $\partial t_{ij}/\partial t$ at a point P in a reference frame such that $v_i = w_{jk} = 0$ at P . We also use the equation of continuity in the equivalent forms

$$\frac{d\rho}{dt} + \rho d_k k = 0, \quad \rho dV = \rho_0 dV_0, \quad (5.2)$$

ρ and V being the density and volume occupied at time t by a set of particles initially occupying the volume V_0 and having density ρ_0 . We require that (5.1) may be inverted to give

$$d_{ij} = B_{ijkm} \frac{Dt_{km}}{dt} \quad (5.3)$$

in some region \mathfrak{R} of stress values, B_{ijkm} being a tensor invariant of t_{pq} with the same symmetry as A_{ijkm} . Here B_{ijkm} , which is the inverse of A_{ijkm} , considered as a 6×6 matrix,

$$A_{ijkm} B_{kmrs} = B_{ijkm} A_{kmrs} = \frac{1}{2} (\delta_{ir} \delta_{js} + \delta_{is} \delta_{jr}), \quad (5.4)$$

exists if and only if $Dt_{ij}/dt = 0$ implies $d_{km} = 0$. Since B_{ijkm} has A_{ijkm} as an inverse,

$$B_{ijkm} C_{ij} = 0, \quad C_{ij} = C_{ji} \text{ implies } C_{ij} = 0. \quad (5.5)$$

In \mathfrak{R} , B_{ijkm} or, equivalently, A_{ijkm} is assumed to be a continuously differentiable function of t_{pq} . We henceforth exclude stress values not in \mathfrak{R} . Continuity requirements for the stress and motion should be evident from the context.

We find it convenient to introduce as independent variables the time and material coordinates X_k such that the motions considered are represented by relations of the form

$$x_i = x_i(X_k, t), \quad X_k = X_k(x_i, t), \quad x_i(X_k, 0) = X_i. \quad (5.6)$$

Then a particle is represented by fixed values of the X_k 's and

$$\frac{d}{dt} = \frac{\partial}{\partial t} \Big|_{X_k}, \quad v_i = \frac{dx_i}{dt}.$$

It should be noted that, under the above assumptions, (5.1) and (5.2) impose no conditions on how the stress depends on the time at a single particle: we can always choose a stress depending only on the time which coincides with any given stress at a given particle at all times, arbitrarily assign functions $w_{ij} = -w_{ji}$ of the time only and a positive constant ρ_0 , use (5.3) to calculate d_{ij} , (5.2) to calculate ρ , and take $v_i = (w_{ij} + d_{ij}) x_j$. With this assignment, (5.1) and (5.2) are satisfied.

6. Work and density functionals

The work $W(t)$ done by the stress in deforming a material volume $V(t)$ during the time interval $\langle 0, t \rangle$ is defined by

$$W(t) \equiv \int_0^t \int_{V(\tau)} t_{ij} d_{ij} dV(\tau) d\tau. \quad (6.1)$$

It is noted in [6] that

$$t_{ij} d_{ij} = F_{ij} \frac{dt_{ij}}{dt} \quad (6.2)$$

whenever (5.1) holds, where

$$F_{ij} = B_{kmij} t_{km} = F_{ji}. \quad (6.3)$$

From (5.2),

$$dV(t) = \frac{\varrho_0}{\varrho} dV_0 = e^0 \int_0^t \dot{a}_{kk} d\tau dV_0, \quad (6.4)$$

where the integral is a partial integration with respect to t holding X_k constant. From (5.3),

$$d_{kk} = G_{ij} \frac{Dt_{ij}}{dt}, \quad (6.5)$$

where

$$G_{ij} = B_{k k i j} = G_{j i} \quad (6.6)$$

is an isotropic function of t_{rs} . As such, it is expressible* in the form

$$G_{ij} = G_0 \delta_{ij} + G_1 t_{ij} + G_2 t_{ik} t_{kj}.$$

Making this substitution in (6.5) and using (5.1), we see that

$$d_{kk} = G_{ij} \frac{dt_{ij}}{dt}, \quad (6.7)$$

the terms involving w_{ij} dropping out in the summation. From (6.4) and (6.5)

$$\frac{\varrho_0}{\varrho} = e^0 \int_0^t G_{ij} dt_{ij}(\tau), \quad (6.8)$$

which expresses the density at X_k as a function of its initial value ϱ_0 and a functional of the stress at X_k . We also have from (6.1), (6.2), (6.4) and (6.8)

$$W(t) = \int_0^t \int_{V_0} t_{ij} d_{ij} \frac{\varrho_0}{\varrho} dV_0 d\tau = \int_{V_0} w(X_k, t) dV_0, \quad (6.9)$$

where

$$w(X_k, t) = \int_0^t t_{ij} d_{ij} \frac{\varrho_0}{\varrho} d\tau = \int_0^t F_{ij} dt_{ij}(\tau) e^0 \int_0^t G_{km} dt_{km}(\tau). \quad (6.10)$$

We thus have the work per unit initial volume at X_k expressed as a functional of the stress at X_k . The tensors F_{ij} , G_{ij} and t_{ij} , being symmetric, have six independent components. With the definitions

$$\begin{aligned} (y^1, y^2, y^3, y^4, y^5, y^6) &= (t_{11}, t_{22}, t_{33}, \tfrac{1}{2}(t_{12} + t_{21}), \tfrac{1}{2}(t_{23} + t_{32}), \tfrac{1}{2}(t_{13} + t_{31})), \\ (F_1, F_2, F_3, F_4, F_5, F_6) &= (F_{11}, F_{22}, F_{33}, 2F_{12}, 2F_{23}, 2F_{13}), \\ (G_1, G_2, G_3, G_4, G_5, G_6) &= (G_{11}, G_{22}, G_{33}, 2G_{12}, 2G_{23}, 2G_{13}), \end{aligned} \quad (6.11)$$

we can write the functional in the form (2.2). Henceforth $\partial \Xi / \partial t_{ij}$, where Ξ is any function of the stress, shall mean

$$\frac{\partial \Xi}{\partial t_{ij}} \equiv \frac{\partial \Xi}{\partial t_{ij}} \left[\frac{1}{2} (t_{mn} + t_{nm}) \right] = \frac{\partial \Xi}{\partial t_{ji}}, \quad (6.12)$$

* See, e.g. [8, Appendix I], setting $a_{ij}^{(1)} = t_{ij}$, $a_{ij}^{(2)} = 0$.

where the nine quantities t_{ij} are regarded as independent in performing the differentiation. With (6.11) and this convention we have, for example, the equivalence of the equations

$$G_{\alpha} = \frac{\partial A}{\partial y^{\alpha}}, \quad G_{ij} = \frac{\partial A}{\partial t_{ij}}.$$

7. Materials of type I

Any assignment of the functions A_{ijkm} subject to the conditions mentioned in Section 5 serves to define a hypo-elastic material. We call such a material *a material of type I* provided its work functional, given by (6.9), is a functional of type I. This amounts to saying that a hypo-elastic material is of type I provided there exists a stress value t_{ij}^0 such that the work $W(t)$ done by the stress in deforming any material volume $V(t)$ is non-negative whenever $t_{ij} = t_{ij}^0$ at $t = 0$, this being true for all motions and stresses consistent with the constitutive equations.

We make free use of (6.11) and (6.12).

From Theorem 1, a material is of type I if and only if there exists a stress value t_{ij}^0 and a non-negative function $\gamma(t_{ij})$ such that $\gamma(t_{ij}^0) = 0$ and such that, for any $t \geq 0$ and fixed X_k ,

$$w(t) = \gamma[t_{ij}(t)] e^{\int_0^t G_{km} dt_{km}(\tau)} \quad (7.1)$$

whenever $t_{ij}(0) = t_{ij}^0$, $w(t)$ being given by (6.10). From Corollary 1, a material is of type I if and only if there exists a stress value t_{ij}^0 and a non-negative function $\gamma(t_{ij})$ such that

$$\frac{\partial \gamma}{\partial t_{ij}} = -\gamma G_{ij} + F_{ij}, \quad \gamma(t_{ij}^0) = 0, \quad (7.2)$$

and, from Corollary 2,

$$F_{ij}(t_{km}^0) = 0. \quad (7.3)$$

Theorem 5. *For any material of type I, $t_{ij}^0 = 0$,*

$$W(t) = \int_{V(t)} \gamma[t_{ij}(x_k, t)] dV(t) \quad \text{when} \quad t_{ij}(X_k, 0) = 0 \quad (7.4)$$

and, regardless of whether $t_{ij}(X_k, 0) = 0$,

$$t_{ij} d_{ij} = \varrho \frac{d}{dt} (\gamma/\varrho). \quad (7.5)$$

Proof. For any material of type I, we have from (6.3) and (7.3)

$$F_{ij}(t_{km}^0) = B_{kmij}(t_{pq}^0) t_{km}^0 = 0,$$

which, from (5.5), implies that $t_{km}^0 = 0$. From (7.1) and (6.8), for fixed X_k ,

$$w(t) = \gamma[t_{ij}(t)] \frac{\varrho_0}{\varrho} \quad \text{when} \quad t_{ij}(0) = 0. \quad (7.6)$$

Integrating (7.6) over V_0 , then using (5.2), we obtain (7.4). From (6.2), (7.2), (6.7) and (5.2),

$$\begin{aligned} t_{ij} d_{ij} &= F_{ij} \frac{dt_{ij}}{dt} = \frac{\partial \gamma}{\partial t_{ij}} \frac{\partial t_{ij}}{\partial t} + \gamma G_{ij} \frac{dt_{ij}}{dt} \\ &= \frac{d\gamma}{dt} + \gamma d_{kk} = \frac{d\gamma}{dt} - \frac{\gamma}{\varrho} \frac{d\varrho}{dt}, \end{aligned}$$

which establishes (7.5).

Thus γ represents the work per unit present volume done by the stress in deforming the material from an initially unstressed state.

Theorem 6. *A hypo-elastic material is of type I if there exists a function $\gamma(t_{ij})$ such that, whenever (5.1) and (5.2) hold,*

$$t_{ij} d_{ij} = \varrho \frac{d}{dt} (\gamma / \varrho), \quad \gamma(t_{ij}) \geq 0, \quad \gamma(0) = 0. \quad (7.7)$$

Proof. Under these hypotheses, we have, using (6.10)

$$w(t) = \int_0^t t_{ij} d_{ij} \frac{\varrho_0}{\varrho} d\tau = \int_0^t \varrho_0 \frac{d}{d\tau} (\gamma / \varrho) d\tau = \frac{\varrho_0}{\varrho} \gamma[t_{ij}(t)] = 0, \quad (7.8)$$

whenever (5.1) and (5.2) hold and $t_{ij}(0) = 0$, X_k being fixed. Since (5.1) and (5.2) do not restrict the values $t_{ij}(\tau)$ for fixed X_k , the functional $w(t)$ is of type I by Definition 1, (7.7) and (7.8).

Theorem 7. *A necessary and sufficient condition that a hypo-elastic material of type I have the property that there exist, in the neighborhood of each stress value, stress values which cannot be attained by deforming the material subject to the condition*

$$t_{ij} d_{ij} - F_{ij} \frac{dt_{ij}}{dt} = 0$$

is that there exist, at least locally, functions $\Phi(t_{ij})$ and $\Psi(t_{ij})$ such that

$$F_{ij} = \Phi \frac{\partial \Psi}{\partial t_{ij}}, \quad (7.9)$$

$$G_{ij} = \frac{1}{\gamma} \left[\Phi \frac{\partial \Psi}{\partial t_{ij}} - \frac{\partial \gamma}{\partial t_{ij}} \right]. \quad (7.10)$$

Proof. It is shown in [6] that a hypo-elastic material has the indicated inaccessibility property if and only if there exist, at least locally, functions Φ and Ψ of the stress such that (7.9) holds. For a material of type I, (7.9) and (7.10) are equivalent since (7.2) holds.

When there exist functions Φ and Ψ satisfying (7.9), Ψ is called a *hypo-elastic potential*.

8. Materials of type II

A hypo-elastic material is a *material of type II* provided the work functional $w(X_k, t)$ is a functional of type II. This is tantamount to requiring that, for all material volumes and for all stresses and motions consistent with the constitutive

equations, the work $W(t)$ done by the stresses in deforming the material be non-negative whenever the stress path is closed,

$$t_{ij}(X_k, t) = t_{ij}(X_k, 0).$$

From Theorem 2, a material is of type II if and only if there exist functions $\Lambda(t_{ij})$ and $\Theta(t_{ij})$ such that, for any $t \geq 0$ and any stress path

$$\int_0^t G_{ij} dt_{ij}(\tau) = \Lambda[t_{ij}(t)] - \Lambda[t_{ij}(0)], \quad (8.1)$$

$$w(t) = \{\Theta[t_{ij}(t)] - \Theta[t_{ij}(0)]\} e^{-\Lambda[t_{ij}(0)]}, \quad (8.2)$$

X_k being held fixed. From Theorem 3, (8.1) and (8.2) can be replaced by

$$G_{ij} = \partial \Lambda / \partial t_{ij}, \quad (8.3)$$

$$F_{ij} = e^{-\Lambda} \partial \Theta / \partial t_{ij}. \quad (8.4)$$

For any such material, $F_{ij} = \Phi \partial \Psi / \partial t_{ij}$, where $\Phi = e^{-\Lambda}$, $\Psi = \Theta$, so a hypo-elastic potential exists. From Theorems 3 and 4, a material is of type I and of type II if and only if (8.3) and (8.4) hold and Θ takes on an absolute minimum at $t_{ij} = 0$, in which case

$$\gamma(t_{ij}) = e^{-\Lambda(t_{ij})} [\Theta(t_{ij}) - \Theta(0)].$$

Also in this case, (7.9) and (7.10) hold with $\Phi = e^{-\Lambda}$, $\Psi = \Theta$.

Theorem 8. Choose any fixed stress value t_{ij}^0 and any material of type II. There exists an admissible choice of Λ and Θ such that, whenever $t_{ij}(X_k, 0) = t_{ij}^0$, we have

$$W(t) = \int_{V_0} \Theta[t_{ij}(X_k, t)] dV_0, \quad (8.5)$$

$$\varrho = \varrho_0 e^{-\Lambda}, \quad (8.6)$$

$$t_{ij} d_{ij} = \frac{\varrho}{\varrho_0} \frac{d\Theta}{dt}. \quad (8.7)$$

Proof. As noted earlier, Λ and Θ are subject to the transformations (4.11). By proper choice of α and β , we can arrange that $\Theta(t_{ij}^0) = \Lambda(t_{ij}^0) = 0$, where t_{ij}^0 is specified. With this choice, we obtain (8.5) from (8.2) and (6.10) and (8.6) from (6.8) and (8.1) when $t_{ij}(X_k, 0) = t_{ij}^0$. From (6.2), (8.4) and (8.5)

$$t_{ij} d_{ij} = F_{ij} \frac{dt_{ij}}{dt} = e^{-\Lambda} \frac{\partial \Theta}{\partial t_{ij}} \frac{dt_{ij}}{dt} = \frac{\varrho}{\varrho_0} \frac{d\Theta}{dt},$$

which is (8.6).

Thus Θ represents the work per unit initial volume done by the stress in deforming the material from a specified state of stress when Θ and Λ are suitably normalized. Under the same conditions, the work per unit present volume Ω is given by

$$\Omega = \Theta \varrho / \varrho_0 = e^{-\Lambda} \Theta, \quad (8.8)$$

both these work densities being functions of stress.

References

- [1] TRUESDELL, C.: The simplest rate theory of pure elasticity. *Comm. Pure and Appl. Math.* **8**, 123—132 (1955).
- [2] TRUESDELL, C.: Hypo-elasticity. *J. Rat. Mech. Anal.* **4**, 83—133 (1955).
- [3] TRUESDELL, C.: Hypo-elastic shear. *J. Appl. Phys.* **27**, 441—447 (1956).
- [4] THOMAS, T. Y.: Isotropic materials whose deformation and distortion energies are expressible by scalar invariants. *Proc. Nat. Acad. Sci.* **42**, 603—608 (1956).
- [5] THOMAS, T. Y.: Deformation energy and the stress-strain relations for isotropic materials. *J. Math. Phys.* **4**, 335—350 (1957).
- [6] ERICKSEN, J. L.: Hypo-elastic potentials. *Quart. J. Mech. Appl. Math.* **11**, 67—72 (1958).
- [7] CAPRIOLI, L.: Su un criterio per l'esistenza dell'energia di deformazione. *Boll. Un. Mat. Ital.* (3) **10**, 481—483 (1955).
- [8] RIVLIN, R. S., & J. L. ERICKSEN: Stress-deformation relations for isotropic materials. *J. Rat. Mech. Anal.* **4**, 323—425 (1955).
- [9] NOLL, W.: On the continuity of the solid and fluid states. *J. Rat. Mech. Anal.* **4**, 3—81 (1955).
- [10] THOMAS, T. Y.: On the structure of the stress-strain relations. *Proc. Nat. Acad. Sci.* **41**, 716—720 (1955).
- [11] THOMAS, T. Y.: Kinematically preferred coordinate systems. *Proc. Nat. Acad. Sci.* **41**, 762—770 (1955).

U. S. Naval Research Laboratory
Washington, D.C.
and
The Johns Hopkins University
Baltimore, Maryland

(Received April 4, 1958)

Steady Shear Flow of Non-Newtonian Fluids

WILLIAM O. CRIMINALE, JR., J. L. ERICKSEN & G. L. FILBEY, JR.

I. Introduction

In recent years, considerable effort has been expended in attempting to gain an understanding of the behavior of non-Newtonian fluids in shear. MARKOVITZ [1] recently attempted to collect and organize experimental data and to compare it with predictions of existing theories. One difficulty is that conclusions which can be drawn from different types of experiments seem to disagree.

Theoretical work on the steady flow of non-Newtonian fluids has brought to light a new phenomenon. According to the linear theory of viscous fluids, it is always possible for a fluid flowing through a cylindrical tube, to which it adheres, to undergo steady rectilinear motion, each particle moving with constant speed in a straight line parallel to the generators of the cylinder. Analyses made by ERICKSEN [2] and by GREEN & RIVLIN [3] show that, for many ideal fluids described by the REINER [4]-RIVLIN [5] theory, this simple type of motion is possible only for very special shapes of tubes and that, in general, it is replaced by a flow consisting of such a rectilinear motion combined with a secondary flow in cross-sectional planes. Using the more general theory of fluids proposed by RIVLIN & ERICKSEN [6], LANGLOIS [7] has obtained solutions involving secondary flows in tubes and other types of boundary geometries. As has been discussed by RIVLIN [8], similar secondary flows have been observed in the mean flow of fluids through non-circular tubes when the motion is turbulent. To our knowledge, it is not known whether secondary flows occur in the steady flow of any real fluid through tubes.

Our analysis indicates that, if the conclusions which ROBERTS [9] drew from his experiments are correct, the fluids which he observed should be capable of undergoing rectilinear motion through tubes of any shape. MARKOVITZ [1] draws self-consistent conclusions from other experimental data which contradict some of ROBERTS' conclusions. Calculations made in accordance with these conclusions suggest that secondary flows are to be expected in some shapes of tubes. We offer a partial proof that if secondary flows do not occur in elliptical tubes, they will not occur in tubes of any other shape.

II. Equations of Motion

For incompressible fluids, the constitutive equations considered are of the form

$$\mathbf{t} = -p\mathbf{I} + \mathbf{f}(A_1, \dots, A_N), \quad (2.1)$$

where \mathbf{t} is the stress tensor, p is an arbitrary pressure, N is any positive integer, the \mathbf{A} 's are symmetric tensors with rectangular Cartesian components given by

$$\begin{aligned}(A_1)_{ij} &= v_{i,j} + v_{j,i}, \\ (A_{M+1})_{ij} &= \frac{\partial}{\partial t} (A_M)_{ij} + (A_M)_{ij,k} v_k + (A_M)_{ik} v_{k,j} + (A_M)_{jk} v_{k,i}, \\ M &= 1, \dots, N-1,\end{aligned}\quad (2.2)$$

where v_k are the velocity components, and \mathbf{f} denotes any symmetric isotropic function of the \mathbf{A} 's, to be determined by experiment. These quantities must satisfy the incompressibility condition

$$v_{k,k} = 0, \quad (2.3)$$

and the equilibrium equations

$$t_{i,j} - \varphi_{,i} = \varrho \left(\frac{\partial v_i}{\partial t} + v_{i,j} v_j \right), \quad (2.4)$$

where ϱ is the density and φ the potential of the body force per unit volume, assumed conservative.

From (2.2) and (2.3),

$$\text{tr } \mathbf{A}_1 = \text{tr } \mathbf{A}_2 - \text{tr } \mathbf{A}_1^2 = 0. \quad (2.5)$$

Whenever

$$\mathbf{A}_M = 0 \quad \text{for } M > 2 \quad \text{or } N = 2, \quad (2.6)$$

\mathbf{f} reduces to a symmetric isotropic function of \mathbf{A}_1 and \mathbf{A}_2 . As such, it is expressible in the form [6], [10]

$$\begin{aligned}\mathbf{f} &= \alpha_0 \mathbf{I} + \alpha_1 \mathbf{A}_1 + \alpha_2 \mathbf{A}_2 + \alpha_3 \mathbf{A}_1^2 + \alpha_4 \mathbf{A}_2^2 + \alpha_5 (\mathbf{A}_1 \mathbf{A}_2 + \mathbf{A}_2 \mathbf{A}_1) + \\ &+ \alpha_6 (\mathbf{A}_1^2 \mathbf{A}_2 + \mathbf{A}_2 \mathbf{A}_1^2) + \alpha_7 (\mathbf{A}_2^2 \mathbf{A}_1 + \mathbf{A}_1 \mathbf{A}_2^2) + \alpha_8 (\mathbf{A}_1^2 \mathbf{A}_2^2 + \mathbf{A}_2^2 \mathbf{A}_1^2),\end{aligned}\quad (2.7)$$

where

$$\begin{aligned}\alpha_r &= \alpha_r(\text{tr } \mathbf{A}_1, \text{tr } \mathbf{A}_1^2, \text{tr } \mathbf{A}_1^3, \text{tr } \mathbf{A}_2, \text{tr } \mathbf{A}_2^2, \text{tr } \mathbf{A}_2^3, \text{tr } \mathbf{A}_1 \mathbf{A}_2, \\ &\quad \text{tr } \mathbf{A}_1^2 \mathbf{A}_2, \text{tr } \mathbf{A}_1 \mathbf{A}_2^2, \text{tr } \mathbf{A}_1^2 \mathbf{A}_2^2),\end{aligned}\quad (2.8)$$

the α 's being polynomials whenever each component of \mathbf{f} is a polynomial in the components of \mathbf{A}_1 and \mathbf{A}_2 . A representation of the form (2.7) with $\alpha_8 = 0$ always exists. With it, the remaining α 's may not be polynomials if \mathbf{f} is. Absorbing α_0 in the arbitrary pressure p and using (2.5), (2.7) and (2.8), we may write (2.1) in the form

$$\begin{aligned}\mathbf{t} &= -p \mathbf{I} + \alpha_1 \mathbf{A}_1 + \alpha_2 \mathbf{A}_2 + \alpha_3 \mathbf{A}_1^2 + \alpha_4 \mathbf{A}_2^2 + \alpha_5 (\mathbf{A}_1 \mathbf{A}_2 + \mathbf{A}_2 \mathbf{A}_1) + \\ &+ \alpha_6 (\mathbf{A}_1^2 \mathbf{A}_2 + \mathbf{A}_2 \mathbf{A}_1^2) + \alpha_7 (\mathbf{A}_2^2 \mathbf{A}_1 + \mathbf{A}_1 \mathbf{A}_2^2) + \alpha_8 (\mathbf{A}_1^2 \mathbf{A}_2^2 + \mathbf{A}_2^2 \mathbf{A}_1^2),\end{aligned}\quad (2.9)$$

where

$$\begin{aligned}\alpha_r &= \alpha_r(\text{tr } \mathbf{A}_1^2, \text{tr } \mathbf{A}_1^3, \text{tr } \mathbf{A}_2^2, \text{tr } \mathbf{A}_2^3, \text{tr } \mathbf{A}_1 \mathbf{A}_2, \text{tr } \mathbf{A}_1^2 \mathbf{A}_2, \text{tr } \mathbf{A}_1 \mathbf{A}_2^2, \text{tr } \mathbf{A}_1^2 \mathbf{A}_2^2), \\ \Gamma &= 1, \dots, 8.\end{aligned}\quad (2.10)$$

Using these equations in essentially this form, RIVLIN [11] has obtained solutions for simple shearing motion, torsional motion, and combined Couette-Poiseuille

flow. The Reiner-Rivlin theory is obtained by setting $N=1$ in (2.1). For it, (2.9) holds with

$$\begin{aligned}\alpha_1 &= \alpha_1 (\text{tr } \mathbf{A}_1^2, \text{tr } \mathbf{A}_1^3), & \alpha_3 &= \alpha_3 (\text{tr } \mathbf{A}_1^2, \text{tr } \mathbf{A}_1^3), \\ \alpha_2 &= \alpha_4 = \alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = 0.\end{aligned}\quad (2.11)$$

III. Laminar Shear Flows

We call a motion a *laminar shear flow* provided that it is possible to choose rectangular Cartesian coordinates at any point P such that, at P

$$\mathbf{A}_1 = \begin{bmatrix} 0 & \sqrt{X} & 0 \\ \sqrt{X} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2X & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{A}_M = 0 \quad \text{for } M > 2, \quad (3.1)$$

where

$$X = \frac{1}{2} \text{tr } \mathbf{A}_1^2. \quad (3.2)$$

For any laminar shear flow, we have from (2.9) and (2.10)

$$\mathbf{t} = -p\mathbf{I} + \beta_1\mathbf{A}_1 + \beta_2\mathbf{A}_2 + \beta_3(\mathbf{A}_1^2 - \frac{1}{2}\mathbf{A}_2), \quad (3.3)$$

where the β 's are functions of X , given by

$$\begin{aligned}\beta_1 &= \gamma_1 + 2\gamma_5 X + 4\gamma_7 X^2, \\ \beta_2 &= \gamma_2 + \frac{1}{2}\gamma_3 + 2(\gamma_4 + \gamma_6)X + 4\gamma_8 X^2, \\ \beta_3 &= \gamma_3, \\ \gamma_r(X) &= \alpha_r(2X, 0, 4X^2, 8X^3, 0, 2X^2, 0, 4X^3).\end{aligned}\quad (3.4)$$

These relations hold in all coordinate systems. MARKOVITZ & WILLIAMSON [12] have used the constitutive equations in a form very similar to and equivalent to (3.3). RIVLIN's general solutions are all laminar shear flows. For a Reiner-Rivlin fluid, it follows from (2.14) and (3.4) that

$$\beta_2 = 2\beta_3. \quad (3.5)$$

For the more general fluids considered here, β_1 , β_2 and β_3 can be assigned arbitrarily as functions of X . We assume that $\beta_1 \neq 0$. It is easily shown that (3.5) holds if and only if the principal directions of \mathbf{t} and \mathbf{A}_1 coincide.

We are primarily concerned with the rectilinear motions

$$v_1 = f(x_2, x_3), \quad v_2 = v_3 = 0, \quad (3.6)$$

which satisfy (2.3). From (2.2) and (3.6),

$$\mathbf{A}_1 = \begin{bmatrix} 0 & f_{,2} & f_{,3} \\ f_{,2} & 0 & 0 \\ f_{,3} & 0 & 0 \end{bmatrix}, \quad \mathbf{A}_2 = 2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & f_{,2}^2 & f_{,2}f_{,3} \\ 0 & f_{,2}f_{,3} & f_{,3}^2 \end{bmatrix}, \quad \mathbf{A}_M = 0 \quad \text{for } M > 2. \quad (3.7)$$

By a rotation of coordinates in the x_2 - x_3 plane, we can always make $f_{,2} \geq 0$, $f_{,3} = 0$ at a point, so that (3.1) holds. Thus these are laminar shear flows such that (3.3) holds. From (3.2) and (3.7),

$$X = f_{,2}^2 + f_{,3}^2 = f_{,i}f_{,i}. \quad (3.8)$$

We have from (3.7) and (3.3)

$$\mathbf{t} = -p\mathbf{I} + \beta_1 \begin{bmatrix} 0 & f_{,2} & f_{,3} \\ f_{,2} & 0 & 0 \\ f_{,3} & 0 & 0 \end{bmatrix} + 2\beta_2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & f_{,2}^2 & f_{,2}f_{,3} \\ 0 & f_{,2}f_{,3} & f_{,3}^2 \end{bmatrix} + \beta_3 \begin{bmatrix} X & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (3.9)$$

Substituting these stresses in (2.4) we obtain

$$\begin{aligned} P_{,1} &= (\beta_1 f_{,i})_{,i}, \\ P_{,2} &= 2(\beta_2 f_{,2} f_{,i})_{,i}, \\ P_{,3} &= 2(\beta_2 f_{,3} f_{,i})_{,i}, \\ P &= p + \varphi. \end{aligned} \quad (3.10)$$

These equations are of the same form as those analyzed in [2], from which we may conclude that, for a solution to exist, there must exist a constant a and a function $R(f)$ such that

$$a = (\beta_1 f_{,i})_{,i}, \quad (3.11)$$

$$\frac{dR}{df} = 2(\beta_2 f_{,i})_{,i}, \quad (3.12)$$

$$P = p + \varphi = R(f) + \int \beta_2 dX + a x_1. \quad (3.13)$$

For a given fluid, a and $R(f)$ may be different for different flows. If

$$2\beta_2 = k\beta_1, \quad k = \text{const.}, \quad (3.14)$$

every solution of (3.11) clearly satisfies (3.12) with $R(f) = k a f$. If

$$\beta_1 = m/\sqrt{X}, \quad m = \text{const.}, \quad (3.15)$$

it follows from STONE'S work [13] that (3.11) implies that the curves $f = \text{const.}$ are straight lines if $a = 0$, circles of radius $1/|ma|$ if $a \neq 0$. If neither (3.14) nor (3.15) holds, it is shown in [2] that, for each choice of a , there exist solutions of (3.11) which do not satisfy (3.12) for any choice of the function $R(f)$.

We define a tube to be a cylinder, S , with generators parallel to the x_1 -axis whose curve of intersection, \mathcal{C} , with a plane $x_1 = \text{const.}$ is a simple closed curve. For a fluid to be capable of undergoing steady rectilinear motion through the interior of S and of adhering to S , there must exist at least one solution of (3.10) inside \mathcal{C} such that

$$f = 0 \quad \text{on } \mathcal{C}. \quad (3.16)$$

For a Newtonian fluid, $\beta_1 = \text{const.}$, $\beta_2 = \beta_3 = 0$, and there exists a unique solution for any given constant a , whenever \mathcal{C} is sufficiently smooth. In this case, (3.14) holds with $k = 0$. For more general fluids, it is to be expected that, when a solution exists, f will be determined essentially uniquely by (3.11), and that the solution thus obtained will satisfy (3.12) for some choice of $R(f)$. A solution may fail to exist because β_1 is such that (3.11) admits no solution satisfying (3.16). As is easily shown, this is the case for all choices of \mathcal{C} if (3.15) holds.

This is the case for so-called perfectly plastic solids, which are formally included in the theory considered. If this difficulty is not encountered, a solution may fail to exist because the solution(s) determined by (3.11) is(are) not consistent with (3.12). From RIVLIN's analysis of Poiseuille flow [11], it would appear that this latter difficulty rarely, if ever, occurs when \mathcal{C} is a circle. From the remarks made above, it is to be expected that this difficulty will be encountered for many choices of \mathcal{C} if (3.14) does not hold for the fluid considered. The analyses made in [3] and [7] indicate that the assumed rectilinear motion is then replaced by a motion of the type

$$\begin{aligned} v_1 &= f(x_2, x_3), & v_2 &= \psi_3, & v_3 &= -\psi_2, \\ \psi &= \psi(x_2, x_3), \end{aligned} \quad (3.17)$$

satisfying the adherence condition

$$v_i = 0 \quad \text{on} \quad \mathcal{C}. \quad (3.18)$$

IV. Experimental Indications

ROBERTS [9] concluded from his experiments that, when

$$v_1 = k x_2, \quad v_2 = v_3 = 0, \quad k = \text{const.} \quad (4.1)$$

the relations

$$t_{22} = t_{33}, \quad t_{11} \neq t_{22} \quad (4.2)$$

hold. From (3.6), (3.9), (4.1) and (4.2),

$$\beta_2 = 0, \quad \beta_3 \neq 0. \quad (4.3)$$

Hence (3.14) should be satisfied with $k=0$. Assuming β_1 is such that (3.11) admits suitable solutions, we see from the discussion of Section III that the fluids which he observed should be capable of undergoing rectilinear motion through tubes of arbitrary shape, to which they adhere.

From data obtained from different types of measurements, MARKOVITZ [1] concluded that (4.2), hence (4.3), does not hold. The quantities ν_1 , ν_2 and $\dot{\gamma}$ which he uses are related to the quantities β_2 , β_3 and X used here by

$$\begin{aligned} 3\nu_1 &= (2\beta_2 + \beta_3) X, \\ 3\nu_2 &= 2(\beta_2 - \beta_3) X, \\ \dot{\gamma} &= |X|. \end{aligned} \quad (4.4)$$

As he mentions, data obtained from torsional shearing of fluids between parallel circular plates indicates that, for a wide range of values of $\dot{\gamma}$, his equation (26) holds quite accurately. It is equivalent to

$$\beta_3 \sqrt{X} + 2 \frac{d}{d\sqrt{X}} \beta_2 X = \alpha, \quad (4.5)$$

where α is a constant. As is indicated by the data summarized in MARKOVITZ' Figure 5, the normal stresses measured in torsional shearing and coaxial cylinder instruments are equal, within experimental error, which means that the left side

of MARKOVITZ' equation (18) should be equal to $-3(\nu_1 + 2\nu_2)$. Translating this equation into our notation, we obtain

$$3(\beta_3 - 2\beta_2)\sqrt{X} = \beta_3\sqrt{X} + 2\frac{d}{d\sqrt{X}}\beta_2 X. \quad (4.6)$$

From (4.5) and (4.6),

$$\begin{aligned} \beta_2\sqrt{X} &= \frac{\alpha}{6} + \frac{\beta}{X}, \\ \beta_2\sqrt{X} &= \frac{2\alpha}{3} + \frac{2\beta}{X}, \end{aligned} \quad (4.7)$$

where β is a constant. From a more detailed and somewhat different analysis of existing data, MARKOVITZ concludes that $\nu_1 = -\nu_2$, which is in agreement with (4.7) if and only if $\beta = 0$. This suggests that

$$\beta_2\sqrt{X} = \frac{\alpha}{6}, \quad \beta_3\sqrt{X} = \frac{2\alpha}{3}. \quad (4.8)$$

in contradiction with (4.3). If (3.14) held, we should have, for some constant k ,

$$\beta_1\sqrt{X} = \frac{\alpha}{3k}. \quad (4.9)$$

Comparison of (4.8) with experimental data indicates that, if we choose units such that $\sqrt{X} = \dot{\gamma}_k$, where $\dot{\gamma}_k$ is MARKOVITZ' "reduced rate of shear", we obtain good agreement in the range $10^3 \leq \sqrt{X} \leq 10^4$. Within this range, the rather meagre, scattered data indicates that

$$\ln \beta_2 \sqrt{X} \big|_{\sqrt{X}=10^4} - \ln \beta_2 \sqrt{X} \big|_{\sqrt{X}=10^3} \ll 1.$$

If (3.14) held, this statement should hold with β_2 replaced by β_1 . According to the viscosity data used by MARKOVITZ,

$$\frac{d \ln \beta_1 \sqrt{X}}{d \ln \sqrt{X}} = 1.43 - 0.12 \ln \sqrt{X},$$

from which we find that

$$\ln \beta_1 \sqrt{X} \big|_{\sqrt{X}=10^4} - \ln \beta_1 \sqrt{X} \big|_{\sqrt{X}=10^3} = 1.01$$

approximately. It would thus seem that (3.14) does not hold, hence that the fluids observed are incapable of undergoing rectilinear motion through tubes of arbitrary shape.

From (3.5) we see that neither (4.3) nor (4.8) is consistent with the Reiner-Rivlin theory. Experimentally, neither (4.3) nor (4.8) has been established conclusively and it is uncertain why contradictory conclusions derive from different experiments. At least until this discrepancy is resolved, it would be premature to conclude the attractively simple Reiner-Rivlin theory to be inadequate. The theory considered in this paper is not as general as that proposed by GREEN & RIVLIN [14] and is, in some respects, less general than that of NOLL [15].

The experimental data referred to above is mainly for polyisobutylene solutions and may not be representative for other types of fluids.

V. Flow Through Elliptical Tubes

In view of the impossibility of conducting experiments using all possible shapes of tubes, it is desirable to have some assurance that, if secondary flows are not observed in tubes of some particular shape, they will not be observed in tubes of any other shape. It seems probable that (3.14) must be satisfied for rectilinear motion to be possible in an elliptical tube. This can be proved for the class of fluids for which $\beta_1 = \mu = \text{const.}$ For these, there is a unique solution of (3.11) which vanishes on the ellipse

$$x_3^2 + p^2 x_2^2 = q^2,$$

where p and q are non-zero constants with $p^2 \neq 1$. It is given by

$$f = r(x_2^2 + p^2 x_3^2 - q^2), \quad r = \frac{1}{2} \frac{a}{\mu} (1 + p^2).$$

For it,

$$\begin{aligned} X &= f_{,i} f_{,i} = 4r^2(x_2^2 + p^4 x_3^2), \\ X_{,i} f_{,i} &= 2f_{,ij} f_{,i} f_{,j} = 16r^3(x_2^2 + p^6 x_3^2) \\ &= 4r[(1 + p^2)X - 4rp^2(f + rq^2)], \\ f_{,ii} &= 2r(1 + p^2). \end{aligned}$$

For (3.12) to be satisfied, we must have

$$\begin{aligned} \frac{dR}{df} &= 2(\beta_2 f_{,i})_{,i} = 2X_{,i} f_{,i} d\beta_2/dX + 2\beta_2 f_{,ii} \\ &= 8r[(1 + p^2)X - 4rp^2(f + rq^2)] d\beta_2/dX + 4\beta_2 r(1 + p^2). \end{aligned} \quad (5.1)$$

Assuming that $r \neq 0$, f and X are functionally independent. We thus have

$$\frac{d^2 R}{df^2} = -32r^2 p^2 d\beta_2/dX,$$

from which

$$\frac{dR}{df} = -32r^2 p^2 \delta f + t, \quad \beta_2 = \delta X + \eta,$$

where δ , t and η are constants. Making these substitutions in (5.1), we obtain

$$t = 12r(1 + p^2) \delta X - 32r^3 p^2 q^2 \delta + 4\eta r(1 + p^2),$$

which is impossible unless $\delta = 0$, in which case $\beta_2 = \eta = \text{const.}$ Thus (3.14) must hold for rectilinear motion to be possible. Solutions illustrating secondary flows in elliptical tubes are given in [3] and [7].

To our knowledge, there is no very conclusive evidence that non-Newtonian fluids adhere to solid boundaries. For fluids such that (3.14) does not hold, it would be reasonable to expect secondary flows to be observed even if some slippage occurs. The effect of slippage should perhaps be considered in deducing the functional forms of the β 's from experimental data, this being one possible source of the discrepancies mentioned earlier.

It is not completely certain that secondary flows will not be observed in the steady flow of fluids through tubes when (3.14) is satisfied. The possibility exists that rectilinear motion and some more complicated type of flow are both possible in some fluids and that these fluids may prefer the more complicated pattern.

References

- [1] MARKOVITZ, H.: Normal stress effect in polyisobutylene solutions. II. Classification and application of rheological theories. *Trans. Soc. Rheol.* **1**, 37—52 (1957).
- [2] ERICKSEN, J. L.: Overdetermination of the speed in rectilinear motion of non-Newtonian fluids. *Quart. Appl. Math.* **14**, 319—321 (1956).
- [3] GREEN, A. E., & R. S. RIVLIN: Steady flow of non-Newtonian fluids through tubes. *Quart. Appl. Math.* **14**, 229—308 (1956).
- [4] REINER, M.: A mathematical theory of dilatancy. *Amer. J. Math.* **67**, 350—362 (1945).
- [5] RIVLIN, R. S.: Hydrodynamics of non-Newtonian fluids. *Nature* **160**, 611—613 (1947).
- [6] RIVLIN, R. S., & J. L. ERICKSEN: Stress-deformation relations for isotropic materials. *J. Rat'l. Mech. Anal.* **4**, 323—425 (1955).
- [7] LANGLOIS, W.: Doctoral Dissertation. Grad. Div. of Applied Math., Brown University 1957.
- [8] RIVLIN, R. S.: The relation between the flow of non-Newtonian fluids and turbulent Newtonian fluids. *Quart. Appl. Math.* **15**, 212—215 (1957).
- [9] ROBERTS, J. E.: Pressure distribution in liquids in laminar shearing motion and comparison with predictions from various theories. *Proc. Second. Inter. Cong. Rheol.* 91—98 (1953).
- [10] RIVLIN, R. S.: Further remarks on the stress-deformation relations for isotropic materials. *J. Rat'l. Mech. Anal.* **4**, 681—702 (1955).
- [11] RIVLIN, R. S.: Solution of some problems in the exact theory of visco-elasticity. *J. Rat'l. Mech. Anal.* **5**, 179—187 (1956).
- [12] MARKOVITZ, H., & R. B. WILLIAMSON: Normal stress effect in polyisobutylene solutions I. Measurements in a cone and plate instrument. *Trans. Soc. Rheol.* **1**, 25—36 (1957).
- [13] STONE, D. E.: On nonexistence of rectilinear motion in plastic solids and non-Newtonian fluids. *Quart. Appl. Math.* **15**, 257—262 (1957).
- [14] GREEN, A. E., & R. S. RIVLIN: The mechanics of non-linear materials with memory. *Arch. Rat'l. Mech. Anal.* **1**, 1—21 (1957).
- [15] NOLL, W.: On the continuity of the solid and fluid states. *J. Rat'l. Mech. Anal.* **4**, 3—81 (1955).

The Johns Hopkins University
Baltimore, Maryland

(Received March 18, 1958)

Asymptotic Solutions

of $u^{iv} + \lambda^2(z u'' + \alpha u' + \beta u) = 0$ for Large $|\lambda|$

ALBERT L. RABENSTEIN

Communicated by C. C. LIN

1. Introduction

The asymptotic theory for the Orr-Sommerfeld equation of hydrodynamic stability has a large literature. The structure of the equation is such that it possesses a *turning point*; and herein lies the main difficulty in the development of such a theory. The equation is of the form

$$(1.1) \quad \varphi^{iv} + \lambda^2 \sum_{n=0}^2 P_n(y) \varphi^{(2-n)} + \sum_{n=0}^2 Q_n(y, \lambda) \varphi^{(2-n)} = 0$$

where the complex parameter λ is supposed to be large in magnitude, the functions $P_n(y)$ are analytic with

$$(1.1a) \quad P_0(y_0) = 0, \quad P'(y_0) \neq 0$$

and

$$(1.1b) \quad Q_n(y, \lambda) \sim \sum_{\nu=0}^{\infty} \frac{q_{n\nu}(y)}{\lambda^{\nu}},$$

where the functions $q_{n\nu}(y)$ are analytic. In the hydrodynamical case, $P_1(y) \equiv 0$. We are interested in the behavior of solutions of (1.1) in a neighborhood of the turning point y_0 when $|\lambda|$ is large.

In order to use the method of a comparison equation to obtain such an asymptotic theory, one needs an equation which is sufficiently similar to (1.1) but which can still be solved explicitly. Such an equation is

$$(1.2) \quad u^{iv} + \lambda^2(z u'' + \alpha u' + \beta u) = 0.$$

Here z is the complex independent variable and α and β are complex parameters, dependent on λ in such a way that they are uniformly bounded for $|\lambda| \geq |\lambda_0|$ for some λ_0 . This equation has a turning point at $z=0$, and is indeed a special case of (1.1). Furthermore it can be solved using the method of Laplace integrals, since the independent variable appears in only one of the coefficients and appears there linearly.

The complete theory for (1.1), based on the functions defined by (1.2), will be presented later*.

* Preliminary forms of this theory have been reported by LIN. See the Bulletin of the Amer. Math. Soc. **62**, 394 (1956), and the Proceedings of the IX^e Congrès International de Mécanique Appliquée, Université de Bruxelles **1**, 137 (1957).

The present paper concerns the behavior of the functions defined by (1.2), for large $|\lambda|$. This equation is clearly a generalization of that considered by WASOW [1]. It turns out that it is necessary to consider (1.2) with general α to make the general theory for (1.1) feasible, even in the case where $P_1(y) \equiv 0$. Also, there are certain general relations that hold for the functions $u(z; \lambda, \alpha, \beta)$, and these are not evident for the functions $u(z; \lambda, 0, \beta)$ considered by WASOW. Indeed, one of the important differences between our theory for (1.1) and those of WASOW [2] and LANGER [3] is the use of the comparison equation (1.2). The present theory follows more closely the earlier work of LANGER for the second order equation [4]; it is simpler in structure and more readily applicable.

Series which satisfy (1.2) formally can be found by well known methods. These have the forms

$$(1.3) \quad \begin{aligned} & \sum_{n=0}^{\infty} \sigma_{1n}(z) \lambda^{-2n}, & \sum_{n=0}^{\infty} \sigma_{2n}(z) \lambda^{-2n}, \\ & e^{\frac{2}{3}i\lambda z^{\frac{3}{2}}} \sum_{n=0}^{\infty} \sigma_{3n}(z) \lambda^{-n}, & e^{-\frac{2}{3}i\lambda z^{\frac{3}{2}}} \sum_{n=0}^{\infty} \sigma_{4n}(z) \lambda^{-n}. \end{aligned}$$

In these, $\sigma_{1n}(z)$ is an entire function, $\sigma_{20}(z)$ has a branch point at $z=0$, and $\sigma_{30}(z) = \sigma_{40}(z) = z^{\frac{\alpha}{2} - \frac{5}{4}}$. Since all solutions of (1.2) are entire functions of z , the multi-valued nature of the expressions (1.3) precludes them from representing actual solutions of (1.2) in a full neighborhood of $z=0$. It is known [5] that one of these formal series may represent one solution in a certain region of the z plane and represent another solution in another region. In any case, the representation is known to be valid only for $|z| \geq |z_0| > 0$, where z_0 is a fixed number, so the immediate vicinity of the turning point is not included in the region of representation.

If we put $\eta = \lambda^{\frac{2}{3}}z$, $u(z) = \chi(\eta)$ in (1.2) we obtain the equation

$$(1.4) \quad \chi^{iv} + \eta \chi'' + \alpha \chi' + \beta \lambda^{-\frac{2}{3}} \chi = 0.$$

Solutions of (1.4) having the form of power series in $\lambda^{-\frac{2}{3}}$ will be valid for $|\eta| < |\eta_0|$ or $|z| < |\lambda|^{-\frac{2}{3}}|\eta_0|$ for an arbitrary but finite η_0 . Thus, for any solution of (1.4), the region of validity in the z plane becomes vanishingly small as $|\lambda|$ becomes large, while the representation of a solution by one of the expressions (1.3) still holds only for $|z| \geq |z_0|$. Also there remains the problem of connecting the solutions of (1.4) with the solutions represented by the expressions (1.3).

In this article we obtain solutions of (1.2) by Laplace integrals. Certain relations which hold among the solutions defined by these integrals are deduced with the aid of CAUCHY's theorem. These relations enable one to describe fully the asymptotic behavior of each particular solution in a full neighborhood of $z=0$.

2. General formulae

Seeking solutions of (1.2) in the form

$$u = \int_C e^{\pi t} f(t) dt$$

where t is a complex variable, we find that f satisfies the first order equation

$$(t^2 f)' = (\lambda^{-2} t^2 + \beta t^{-2} + \alpha t^{-1}) (t^2 f)$$

where f and C are to be such that $[e^{zt} t^2 f]_C = 0^*$. Then a solution of (1.2) is

$$(2.1) \quad u = \int_C t^{\alpha-2} \exp\left(\frac{1}{3} \lambda^{-2} t^3 - \beta t^{-1} + z t\right) dt,$$

where C is to be such that

$$(2.2) \quad [t^\alpha \exp\left(\frac{1}{3} \lambda^{-2} t^3 - \beta t^{-1} + z t\right)]_C = 0.$$

If repeated differentiation under the integral sign is permissible, there follows from (2.1) that

$$(2.3) \quad \frac{d^n u(z, \alpha)}{dz^n} = u(z, \alpha + n).$$

It will be seen that this is the case for every solution considered here. Thus, if the asymptotic behavior of a solution (2.1) can be obtained for general α , the behavior of its derivatives follows from (2.3). From (1.2) and (2.3) we also obtain the recurrence formula

$$(2.4) \quad u(z, \alpha + 4) = -\lambda^2 [z u(z, \alpha + 2) + \alpha u(z, \alpha + 1) + \beta u(z, \alpha)].$$

In what follows, we suppose λ to be large in magnitude and to have constant argument. Having chosen a particular value for $\arg \lambda$, by λ^p we shall mean

$$\lambda^p = \exp\{p(\log |\lambda| + i \arg \lambda)\}$$

for all complex p . We assume $\beta \neq 0$, but α and β are otherwise arbitrary. As remarked in the introduction, they may depend on λ in such a way that they are uniformly bounded for $|\lambda| \geq |\lambda_0|$. We take for $\arg \beta$ the value which lies in the range

$$(2.5) \quad \frac{2}{3} \arg \lambda \leq \arg \beta < 2\pi + \frac{2}{3} \arg \lambda.$$

We now define particular solutions of (1.2) by making appropriate choices for the path C of integration in (2.1). Because of the essential singularity of the integrand at $t=0$, paths tending to the origin must do so with direction $(\arg \beta + \gamma \frac{\pi}{2}) \pmod{2\pi}$, where $-1 < \gamma < 1$. Because of the factor t^α , the integrand in (2.1) is in general multiple-valued. The quantity $\lambda^{-2} t^3$ has a negative real part in the sectors

$$\begin{aligned} T_1: & \quad \frac{\pi}{6} + \frac{2}{3} \arg \lambda < \arg t < \frac{\pi}{2} + \frac{2}{3} \arg \lambda \\ T_2: & \quad \frac{5\pi}{6} + \frac{2}{3} \arg \lambda < \arg t < \frac{7\pi}{6} + \frac{2}{3} \arg \lambda \\ T_3: & \quad -\frac{\pi}{2} + \frac{2}{3} \arg \lambda < \arg t < -\frac{\pi}{6} + \frac{2}{3} \arg \lambda. \end{aligned}$$

These are the unshaded sectors in Figure 1. We also define

$$\bar{T}_2: \quad -\frac{7\pi}{6} + \frac{2}{3} \arg \lambda < \arg t < -\frac{5\pi}{6} + \frac{2}{3} \arg \lambda.$$

* For a discussion of the use of the Laplace integral in solving differential equations, see CODDINGTON & LEVINSON, *Theory of Ordinary Differential Equations*, New York, 1955, p. 170.

In Figure 1, T_2 and \bar{T}_2 have the same location, but the function t^α will have in general* different values in T_2 and \bar{T}_2 . We define three solutions $u = A_k$, $k = 1, 2, 3$, by taking paths $C(A_k)$ going from infinity to infinity in two unshaded sectors. The path $C(A_1)$ goes to infinity in \bar{T}_2 and T_3 , the path $C(A_2)$ goes to infinity in T_3 and T_1 , and the path $C(A_3)$ goes to infinity in T_1 and T_2 .

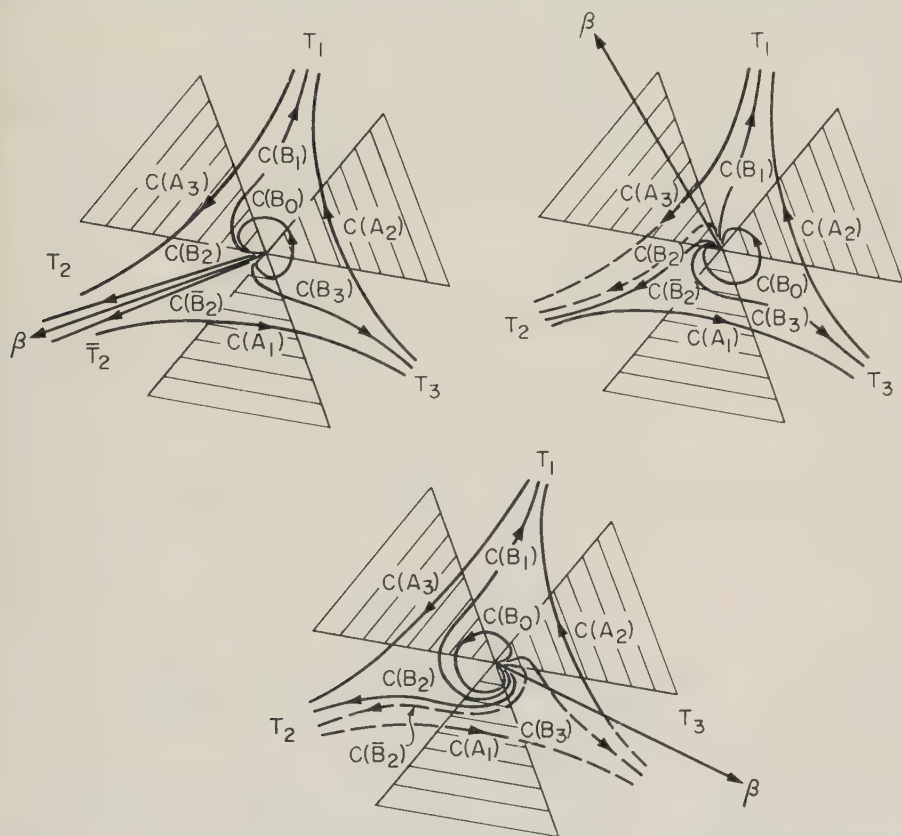


Fig. 1. Typical configurations in the t -plane

We obtain three more solutions $u = B_k$, $k = 1, 2, 3$, by taking paths from the origin to infinity. The paths $C(B_k)$ are described as follows. The path $C(B_1)$ goes to infinity in the sector T_1 and to zero with direction $\arg t = \arg \beta$. The path $C(B_2)$ goes to infinity in T_2 and to zero with direction $\arg t = \arg \beta$. The path $C(B_3)$ goes to infinity in T_3 and goes to zero with direction $\arg t = \arg \beta - 2\pi$.

It will be convenient to define a solution $\bar{B}_2 = e^{-2\pi i \alpha} B_2$ by taking a path $C(\bar{B}_2)$ going to infinity in the sector \bar{T}_2 and tending to the origin with direction $\arg t = \arg \beta - 2\pi$.

We also define a solution $u = B_0$ corresponding to a path $C(B_0)$ which emerges from the origin with direction $\arg t = \arg \beta - 2\pi$, circles the origin, and tends to zero again with direction $\arg t = \arg \beta$.

* In other words, T_2 and \bar{T}_2 will lie on different sheets of the Riemann surface for the function $t^\alpha = \exp[\alpha(\log |t| + i \arg t)]$.

Convergence at the origin is not altered if the direction of approach is changed by an angle less than $\frac{1}{2}\pi$, in accordance with the remarks following (2.5). Use will be made of this later. For each path C described above, the integral (2.1) converges uniformly for z in any finite neighborhood S of $z=0$ and for α in any bounded region of the α plane. Also, the condition (2.2) is satisfied. Then differentiation under the integral sign is valid, and (2.3) follows. These paths of integration are shown in Figure 1 for various values of $\arg \beta$. In the visible portion of the Reimann surface, $\arg t$ ranges from $\arg \beta - 2\pi$ to $\arg \beta$. The dotted parts of the paths are the parts which lie on other sheets of the surface. The paths which go to infinity do so with directions independent of β . At the origin, the direction of approach varies continuously with $\arg \beta$, for $\arg \beta$ in the range (2.5). As a consequence of CAUCHY'S theorem, the following relations hold among the solutions:

$$(2.6) \quad \begin{aligned} \bar{B}_2 + A_1 &= B_3, \\ B_3 + A_2 &= B_1 + B_0 \quad (\bar{B}_2 = B_2 e^{-2\pi i \alpha}), \\ B_1 + A_3 &= B_2. \end{aligned}$$

The paths of integration have been chosen so that these relations hold for all values of β .

We now seek the asymptotic description of each of the functions defined above in a full neighborhood of the turning point $z=0$. The development follows closely that used by WASOW in [1].

3. The solution B_0

We have

$$(3.1) \quad B_0 = \int_{C(B_0)} t^{\alpha-2} \exp(zt - \beta t^{-1}) \exp\left(\frac{1}{3} \lambda^{-2} t^3\right) dt.$$

Since the series

$$\exp\left(\frac{1}{3} \lambda^{-2} t^3\right) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{1}{3} \lambda^{-2} t^3\right)^n$$

converges uniformly for t in any bounded domain and $|\lambda| \geq |\lambda_0| > 0$, and since the rest of the integrand in (3.1) is uniformly bounded for t on $C(B_0)$, termwise integration is permissible and we obtain

$$(3.2) \quad B_0 = b_0(z, \alpha) + \frac{1}{3} \lambda^{-2} b_0(z, \alpha + 3) + \dots$$

where

$$(3.3) \quad b_0(z, \alpha) = \int_{C(B_0)} t^{\alpha-2} \exp(zt - \beta t^{-1}) dt.$$

This may be expressed in terms of a Bessel function. The change of variable $t = e^{-\pi i} \beta \tau^{-1}$ gives

$$b_0(z, \alpha) = e^{-\pi i \alpha} \beta^{\alpha-1} \int_{C'(B_0)} \tau^{-\alpha} \exp\left(\tau - \frac{\beta z}{\tau}\right) d\tau.$$

Comparison with the integral representation formulas for a Bessel function of the first kind ([6], p. 176) gives

$$(3.4) \quad b_0(z, \alpha) = -2\pi i e^{-\pi i \alpha} (z^{\frac{1}{2}} \beta^{-\frac{1}{2}})^{1-\alpha} J_{\alpha-1}(2z^{\frac{1}{2}} \beta^{\frac{1}{2}}).$$

This may be written

$$(3.5) \quad b_0(z, \alpha) = -2\pi i e^{-\pi i \alpha} \beta^{\alpha-1} \sum_{n=0}^{\infty} \frac{(-1)^n (\beta z)^n}{n! \Gamma(n + \alpha)}.$$

Thus B_0 tends to a single-valued analytic function of z in a full neighborhood of $z=0$. We note here that $b_0(z, \alpha)$ satisfies the second order equation

$$(3.6) \quad z u'' + \alpha u' + \beta u = 0$$

obtained from (2.1) by formally allowing λ to become infinite.

4. The solutions A_k

We have

$$(4.1) \quad A_k = \int_{C(A_k)} t^{\alpha-2} \exp\left(\frac{1}{3} \lambda^{-2} t^3 + z t - \beta t^{-1}\right) dt.$$

We shall evaluate A_k asymptotically by the method of steepest descents ([7], p. 503). For this purpose let us write

$$(4.2) \quad t = \nu \tau$$

and

$$\frac{1}{3} \lambda^{-2} t^3 + z t - \frac{1}{3} \lambda^{-2} \nu^3 \tau^3 + \nu z \tau - \mu \left(\frac{1}{3} \tau^3 - \tau\right).$$

Then

$$(4.3) \quad \mu = \lambda^{-2} \nu^3 = -\nu z,$$

whence

$$(4.4) \quad \nu = -i \lambda z^{\frac{1}{2}}, \quad \mu = i \lambda (z^{\frac{1}{2}})^3,$$

where the same branch of the square root is to be taken. From the formal solutions (1.3) we expect the quantity

$$(4.5) \quad \xi = \frac{2}{3} i \lambda z^{\frac{3}{2}} = \frac{2}{3} \mu$$

to be important. In terms of ξ , (4.3) gives

$$(4.6) \quad \nu = \lambda^{\frac{3}{2}} \left(\frac{3}{2} \xi\right)^{\frac{1}{2}}, \quad z = -\lambda^{-\frac{3}{2}} \left[\left(\frac{3}{2} \xi\right)^{\frac{3}{2}}\right]^2,$$

where the same branch of the cube root is taken. The proper branches will be determined in the sequel. Let us now write

$$\frac{1}{3} \lambda^{-2} t^3 + z t = \xi \left(\frac{1}{2} \tau^3 - \frac{3}{2} \tau\right) = |\xi| f(\tau)$$

where

$$(4.7) \quad f(\tau) = e^{i\vartheta} \left(\frac{1}{2} \tau^3 - \frac{3}{2} \tau\right), \quad \vartheta = \arg \xi.$$

Then (4.1) becomes

$$(4.8) \quad A_k = \left[\lambda^{\frac{3}{2}} \left(\frac{3}{2} \xi\right)^{\frac{1}{2}}\right]^{\alpha-1} \int_{C'(A_k)} F(\tau, \xi, \lambda) e^{|\xi| f(\tau)} d\tau,$$

where

$$F(\tau, \xi, \lambda) = \tau^{\alpha-2} \exp\left\{-\beta \lambda^{-\frac{3}{2}} \left(\frac{3}{2} \xi\right)^{-\frac{1}{2}} \tau^{-1}\right\}.$$

$f(\tau)$ is given by (4.7) and $C'(A_k)$ is the image of $C(A_k)$ in the τ plane. Clearly the location of $C'(A_k)$ depends on the choice of a branch for ν in (4.6).

We consider first the case where the variable ξ is large in magnitude, say $|\xi| \geq |\xi_0|$ for some ξ_0 .

The function $\operatorname{Re}\{f(\tau)\}$ has a saddle point at $\tau=1$ where $f'(\tau)=0$, and

$$(4.9) \quad f(\tau) = e^{i\vartheta} \left[-1 + \frac{3}{2}(\tau-1)^2 + \frac{1}{2}(\tau-1)^3 \right].$$

The path of steepest descent through $\tau=1$ is given by

$$(4.10) \quad \operatorname{Im}\{f(\tau) - f(1)\} = 0.$$

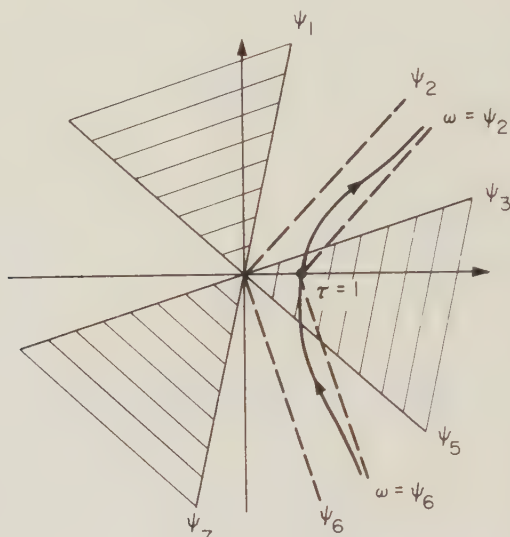


Fig. 2. Path of steepest descent

$$\begin{aligned} & \tau\text{-plane} \\ \psi_j &= \frac{\pi}{2} - (j-1)\frac{\pi}{6} - \frac{\vartheta - 2m\pi}{3}, \quad j = 1, 2, 3, 5, 6, 7 - \pi < \vartheta - 2m\pi < \pi, \quad \psi = \arg \tau \end{aligned}$$

To plot this path, put $\tau-1 = \rho e^{i\omega}$. Then (4.10) becomes

$$\rho^2 [\rho \sin(3\omega + \vartheta) + 3 \sin(2\omega + \vartheta)] = 0.$$

For ϑ in each of the ranges

$$(4.11) \quad -\pi + 2m\pi < \vartheta < \pi + 2m\pi \quad (m = 0, \pm 1, \dots),$$

the path of steepest descent is as shown in Figure 2. In the Figure the unshaded sectors are those in which $e^{i\vartheta}\tau^3$ has a negative real part. For each range of ϑ in (4.11), the path of steepest descent goes to ∞ in the two unshaded sectors indicated. As ϑ passes out of any one of the ranges (4.11), the path changes abruptly and goes to ∞ in the third unshaded sector.

The idea now is to choose, for each A_k , a branch for v in (4.6) so that for ξ in the region specified by (4.11), $C'(A_k)$ goes to infinity in the same two sectors of the τ -plane as does the path of steepest descent. Then we can take for $C'(A_k)$ this path of steepest descent. This requirement is satisfied if we take $m = k-2$ in (4.11) so that

$$(4.12) \quad -\pi + 2\pi(k-2) < \vartheta < \pi + 2\pi(k-2)$$

and take

$$(4.13) \quad \nu = \lambda^{\frac{2}{3}} \left| \frac{3}{2} \xi \right|^{\frac{1}{3}} e^{\frac{1}{3} i \vartheta}$$

for ϑ in the range (4.12). To see that this is so, let $C(A_k)$ go to infinity in the t -plane with the directions

$$\arg t = \pm \frac{\pi}{3} + \frac{2}{3} \arg \lambda + \frac{2\pi}{3} (k-2).$$

Then since $\arg \tau = \arg t - \arg \nu$, we see that $C'(A_k)$ goes to infinity in the τ -plane with directions

$$\arg \tau = \pm \frac{\pi}{3} - \frac{\vartheta - 2\pi(k-2)}{3}.$$

Inspection of Figure 2 with m replaced by $k-2$ shows that the choices (4.12) and (4.13) are satisfactory.

To evaluate A_k by the method of steepest descents, set

$$(4.14) \quad f(\tau) - f(1) = e^{i\vartheta} \left[\frac{3}{2} (\tau-1)^2 + \frac{1}{2} (\tau-1)^3 \right] = -\frac{1}{2} s^2.$$

Choose

$$(4.15) \quad s = (-1)^{k+1} i e^{\frac{1}{2} i \vartheta} (\tau-1) (\tau+2)^{\frac{1}{2}}$$

with the branch cut along the negative real axis from $\tau = -2$ to $\tau = \infty$, and take the branch of the square root which is real and positive for τ real and $\tau > -2$. Then as τ traverses the path of steepest descent in the direction corresponding to that for $C'(A_k)$, s increases through real values. The derivative $ds/d\tau$ is finite and non-zero in the cut plane, so near $s=0$, τ is an analytic function of s . For real s , and $|s| \geq \delta > 0$,

$$|F(\tau(s), \xi, \lambda)| \leq K_1 |s|^{\frac{2}{3}(\alpha-2)}, \quad \left| \frac{d\tau}{ds} \right| \leq K_2.$$

Hence there exist positive constants K and a such that

$$|F(\tau(s), \xi, \lambda)| \cdot \left| \frac{d\tau}{ds} \right| \leq K e^{\frac{1}{2} a s^2}$$

for real s . If we expand $F d\tau/ds$ in a Taylor series about $s=0$, it follows from WATSON'S lemma ([7], p. 501) that the series obtained formally by termwise integration over $(-\infty, \infty)$ represents A_k asymptotically for

$$(4.16) \quad -\frac{3\pi}{2} + 2\pi(k-2) < \vartheta < \frac{3\pi}{2} + 2\pi(k-2).$$

WATSON'S lemma extends the range of ϑ in (4.12) by $\frac{1}{2}\pi$ in each direction. It can be verified from (4.4) that taking ϑ in the range (4.16) and choosing ν as in (4.13) is equivalent to restricting $\varphi = \arg z$ to the range

$$(4.17) \quad \frac{4\pi}{3} (k-2) - \frac{2}{3} \arg \lambda < \varphi < 2\pi + \frac{4\pi}{3} (k-2) - \frac{2}{3} \arg \lambda$$

and taking

$$(4.18) \quad \nu = e^{-\frac{\pi i}{2}} \lambda |z|^{\frac{1}{3}} e^{\frac{1}{3} i \varphi},$$

$$(4.19) \quad \xi = \frac{2}{3} i \lambda |z|^{\frac{2}{3}} e^{\frac{2}{3} i \varphi}.$$

The results of the integration are

$$(4.20) \quad A_k \sim i \sqrt{\pi} (-1)^k \lambda^{\frac{2}{3}(\alpha-1)} \left(\frac{2}{3}\xi\right)^{\frac{1}{3}\alpha-\frac{5}{6}} e^{-\xi} \left[1 + \left(\frac{2}{3}\beta z - (\alpha-3)^2 + \frac{1}{6}\right)\xi^{-1} + \dots\right],$$

where $\vartheta = \arg \xi$ is in the range (4.16). This can be written

$$(4.21) \quad A_k = i \sqrt{\pi} (-1)^k e^{-\frac{3\pi i}{2}(\frac{\alpha}{3} + \frac{1}{2})} \lambda^{\alpha-\frac{3}{2}} z^{\frac{\alpha}{2}-\frac{5}{4}} e^{-\frac{2}{3}i\lambda z^{\frac{2}{3}}} [1 + O(\xi^{-1})]$$

where $\varphi = \arg z$ is in the range (4.17).

Let us consider again the function ξ in (4.5). In the z plane, $\operatorname{Re}(\xi)$ is zero on three rays, denoted by C_k , $k=1, 2, 3$, extending from the origin. These rays are given (mod 2π) by

$$C_k: \quad \arg z = \frac{2\pi}{3} - (k-1)\frac{2}{3}\pi - \frac{2}{3}\arg \lambda.$$

These divide the finite neighborhood S of $z=0$ into three sectors S_k of angle $\frac{2}{3}\pi$. We take S_k to be the sector opposite C_k and take S_k to be closed. It can now be verified that the choice of a branch for ξ in (4.17), (4.19) is equivalent to restricting z to $S - C_k$ and choosing the branch of ξ for which $\operatorname{Re}(\xi) \geq 0$ in S_k and $\operatorname{Re}(\xi) \leq 0$ in the other two sectors.

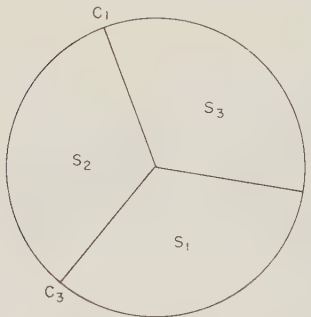


Fig. 3. Diagram for choice of branches

If we examine the relations (4.20) we see that for z in any closed subset of the interior of S_k , A_k tends exponentially to zero as λ becomes infinite and becomes infinite exponentially for z in a closed subset of the interior of either of the other two sectors. The behavior is described by saying that A_k is *subdominant* or *recessive* in S_k and *dominant* in the other two sectors.

The behavior of A_k near C_k can be obtained from the relations (2.6), once the asymptotic properties of the solutions B_k have been investigated.

For the case when $|\xi| < |\xi_0|$ we put $t = \lambda^{\frac{2}{3}}\sigma$, $z = \lambda^{-\frac{2}{3}}\eta$ in (4.1). We note that $|\xi| < |\xi_0|$ implies $|\eta| < |\eta_0|$. These substitutions give

$$(4.22) \quad A_k = \lambda^{\frac{2}{3}(\alpha-1)} \int_{C^*(A_k)} \sigma^{\alpha-2} \exp\left(\frac{\sigma^3}{3} + \eta\sigma\right) \exp(-\beta\lambda^{-\frac{2}{3}}\sigma^{-1}) d\sigma.$$

If we write

$$\exp(-\beta\lambda^{-\frac{2}{3}}\sigma^{-1}) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{\beta\lambda^{-\frac{2}{3}}}{\sigma}\right)^n$$

and integrate term by term, this being valid here, we get

$$(4.23) \quad A_k = \lambda^{\frac{2}{3}(\alpha-1)} \sum_{n=0}^{\infty} \frac{(-\beta\lambda^{-\frac{2}{3}})^n}{n!} g_k(\eta, \alpha-n),$$

where

$$(4.24) \quad g_k(\eta, \alpha) = \int_{C^*(A_k)} \sigma^{\alpha-2} \exp\left(\frac{\sigma^3}{3} + \eta\sigma\right) d\sigma.$$

We observe from (4.24) that

$$(4.25) \quad \frac{d^n g_k(\eta, \alpha)}{d\eta^n} = g_k(\eta, \alpha + n).$$

The functions $g_k(\eta, \alpha)$ satisfy the equation

$$(4.26) \quad \chi^{iv} + \eta \chi'' + \alpha \chi' = 0$$

obtained from (1.4) by putting formally $\lambda = \infty$. Then because of (4.25),

$$(4.27) \quad g_k(\eta, \alpha + 4) + \eta g_k(\eta, \alpha + 2) + \alpha g_k(\eta, \alpha + 1) = 0.$$

This holds for all α , and therefore

$$(4.28) \quad g_k(\eta, \alpha + 3) + \eta g_k(\eta, \alpha + 1) + (\alpha - 1) g_k(\eta, \alpha) = 0.$$

Hence $g_k(\eta, \alpha)$ actually satisfies the third order equation

$$(4.29) \quad \chi''' + \eta \chi' + (\alpha - 1) \chi = 0.$$

For a discussion of this equation see LANGER [8] and HERSHENOV*.

5. The solutions B_k , $k = 1, 2, 3$

We have

$$(5.1) \quad B_k = \int_{C(B_k)} t^{\alpha-2} \exp\left(\frac{1}{3} \lambda^{-2} t^3 + z t - \beta t^{-1}\right) dt.$$

Let

$$(5.2) \quad b_k(z, \alpha) = \int_{C(B_k)} t^{\alpha-2} \exp(tz - \beta t^{-1}) dt.$$

This integral converges for all α if $\operatorname{Re}(tz) < 0$ for large $|t|$. The requirements $\operatorname{Re}(tz) < 0$ and $\operatorname{Re}(\lambda^{-2} t^3) < 0$ for large $|t|$ can be satisfied by deformation of the path $C(B_k)$ if and only if $z \in S - S_k$. Now

$$e^{\frac{1}{3} \lambda^{-2} t^3} = \sum_{j=0}^n \frac{1}{j!} \left(\frac{1}{3} \lambda^{-2} t^3 \right)^j + \lambda^{-2(n+1)} t^{3(n+1)} h_n(t, \lambda),$$

where $h_n(t, \lambda)$ is analytic in t for fixed λ and is uniformly bounded for t on $C(B_k)$ and $|\lambda| \geq |\lambda_0| > 0$. Therefore (5.4) gives

$$(5.3) \quad B_k = b_k(z, \alpha) + \cdots + \frac{1}{n!} \left(\frac{1}{3} \lambda^{-2} \right)^n b_k(z, \alpha + 3n) + R_n,$$

where

$$(5.4) \quad R_n = \lambda^{-2(n+1)} \int_{C(B_k)} t^{\alpha+3(n+1)-2} \exp(tz - \beta t^{-1}) h_n(t, \lambda) dt.$$

If $|z| \geq |z_0| > 0$, then

$$(5.5) \quad R_n = O(\lambda^{-2(n+1)}).$$

In the case where $|\xi| \geq |\xi_0|$ but z is not bounded away from the origin, put $\tau = tz$ in (5.4). Then

$$(5.6) \quad R_n = \lambda^{-2(n+1)} z^{-(\alpha+3n+z)} \int_{C'(B_k)} \tau^{\alpha+3n+1} \exp\left(\tau - \frac{\beta z}{\tau}\right) h_n^*(\tau, z, \lambda) d\tau.$$

* Master's thesis, Department of Mathematics, M.I.T. (1957).

For $n > -\frac{2}{3} - \frac{1}{3} \operatorname{Re}(\alpha)$, the integral here is uniformly bounded and

$$(5.7) \quad R_n = O(z^{1-\alpha} \xi^{-2(n+1)}).$$

Later it will be seen that regardless of the value of α ,

$$(5.8) \quad B_k = b_k(z, \alpha) [1 + O(\xi^{-2})].$$

For some values of α the error term $O(\xi^{-2})$ can be replaced by one of still smaller order of magnitude.

To express the functions $b_k(z, \alpha)$ in terms of special functions, we make the change of variable

$$(5.9) \quad s = t^{\frac{1}{2}} \beta^{-\frac{1}{2}}$$

in (5.2), obtaining

$$b_k = (z^{\frac{1}{2}} \beta^{-\frac{1}{2}})^{1-\alpha} \int_{C^*(B_k)} s^{\alpha-2} \exp \left[z^{\frac{1}{2}} \beta^{\frac{1}{2}} \left(s - \frac{1}{s} \right) \right] ds.$$

Comparison with the integral representation formulae for the Hankel functions (see [6], p. 179) shows that with the proper choice of branches in (5.9),

$$(5.10a) \quad b_1 = \pi i (\beta^{-\frac{1}{2}} z^{\frac{1}{2}})^{1-\alpha} H_{1-\alpha}^{(1)}(2\beta^{\frac{1}{2}} z^{\frac{1}{2}}),$$

$$(5.10b) \quad b_2 = \pi i (\beta^{-\frac{1}{2}} z^{\frac{1}{2}})^{1-\alpha} H_{1-\alpha}^{(1)}(2\beta^{\frac{1}{2}} z^{\frac{1}{2}}),$$

$$(5.10c) \quad b_3 = \pi i (e^{2\pi i} \beta^{-\frac{1}{2}} z^{\frac{1}{2}})^{1-\alpha} H_{1-\alpha}^{(1)}(2\beta^{\frac{1}{2}} z^{\frac{1}{2}})$$

where $z^{\frac{1}{2}} = |z|^{\frac{1}{2}} e^{\frac{1}{2}i\varphi}$,

$$(5.11) \quad -\frac{2\pi}{3} - \frac{2\pi}{3}(k-2) - \frac{2}{3} \arg \lambda < \varphi < \frac{2\pi}{3} - \frac{2\pi}{3}(k-2) - \frac{2}{3} \arg \lambda,$$

and $\arg \beta^{\pm \frac{1}{2}} = \pm \frac{1}{2} \arg \beta$, with $\arg \beta$ as in (2.5). Examination of these functions for various values of α proves (5.8). We remark that the functions $b_k(z, \alpha)$ satisfy the second order equation (3.6).

We now give a more detailed analysis for the case $k=2$, using the formula in [6], p. 179. For large $|t|$ we wish to have

$$\frac{5\pi}{6} + \frac{2}{3} \arg \lambda < \arg t < \frac{7\pi}{6} + \frac{2}{3} \arg \lambda$$

according to the choice for $C(B_2)$, and we want

$$\frac{\pi}{2} + 2n\pi < \arg(tz) < \frac{3\pi}{2} + 2n\pi$$

for some integer n , so that the integral b_2 converges. These will hold if we take

$$\arg t = -\frac{1}{4} \arg z + \pi + \frac{1}{2} \arg \lambda + \frac{1}{2} n\pi,$$

where

$$-\frac{2}{3} \pi + 2n\pi - \frac{2}{3} \arg \lambda < \arg z < \frac{2\pi}{3} + 2n\pi - \frac{2}{3} \arg \lambda,$$

for large $|t|$. For the moment take $\arg \beta^{\frac{1}{2}} = \frac{1}{2} \arg \beta + m\pi$. In the s -plane, the path of integration tends to infinity in the direction

$$\arg s = \pi + \frac{1}{4} \arg z + \frac{1}{2} \arg \lambda + \frac{1}{2} n\pi - m\pi - \frac{1}{2} \arg \beta = \pi - \omega.$$

(See [6], p. 179 for the significance of ω .) The direction of approach for $C(B_2)$ at the origin in the t -plane can be taken as $\arg t = \arg \beta + \gamma \frac{\pi}{2}$, where $-1 < \gamma < 1$. We choose γ so that at the origin in the s -plane

$$\arg s = \frac{1}{2} \arg z + \frac{1}{2} \arg \beta - m\pi + \gamma \frac{\pi}{2} = \omega.$$

The two expressions for ω give

$$\gamma = \frac{\pi}{2} \left(2m\pi - \frac{n\pi}{2} - \frac{1}{2} \arg \lambda - \frac{3}{4} \arg z \right).$$

To have $|\gamma| < 1$ we must have $m - n = 0$. Next, we find

$$-\frac{\pi}{2} - 2n\pi < \omega - \arg(\beta^{\frac{1}{2}} z^{\frac{1}{2}}) < \frac{\pi}{2} - 2n\pi,$$

so we take $n = 0$, and then $m = 0$. Here $-\frac{\pi}{6} < \omega < \frac{7}{6}\pi$. The formula in [6] is given for $|\omega| < \pi$, but extending the range of ω simply gives an analytic continuation of the Hankel function. The relations (5.10b) and (5.11) with $k = 2$ now follow.

The asymptotic behavior of B_k in S_k can be obtained from the relations (2.6). We note that B_k is dominant in S_k . For z in $S - S_k$ and $|z| \geq |z_0| > 0$, B_k tends to a finite non-zero limit as λ becomes infinite. We say that B_k is *balanced* in $S - S_k$.

We now examine the functions B_k in the case where $|\xi| < |\xi_0|$. Here it is impossible to proceed as with the solutions A_k because the path of integration goes to the origin.

When α is not an integer we can solve the equations (2.6) for the three functions B_k in terms of the functions B_0 and A_k . We get

$$(5.12a) \quad B_1 = \frac{A_1 + A_2 - B_0 + e^{-2\pi i \lambda A_3}}{1 - e^{-2\pi i \alpha}},$$

$$(5.12b) \quad B_2 = \frac{A_1 + A_2 + A_3 - B_0}{1 - e^{-2\pi i \alpha}},$$

$$(5.12c) \quad B_3 = \frac{A_1 + e^{-2\pi i \alpha} (A_2 + A_3 - B_0)}{1 - e^{-2\pi i \alpha}}.$$

The asymptotic expansions for the functions B_0 and A_k when $|\xi| < |\xi_0|$ are given by (3.2) and (4.23).

Since all solutions here are entire functions of α , we can pass to the limit as α tends to an integer n , obtaining

$$(5.13a) \quad B_1 = \frac{1}{2\pi i} \left[\frac{\partial}{\partial \alpha} (A_1 + A_2 + A_3 - B_0) - 2\pi i A_3 \right]_{\alpha=n},$$

$$(5.13b) \quad B_2 = \frac{1}{2\pi i} \left[\frac{\partial}{\partial \alpha} (A_1 + A_2 + A_3 - B_0) \right]_{\alpha=n},$$

$$(5.13c) \quad B_3 = \frac{1}{2\pi i} \left[\frac{\partial}{\partial \alpha} (A_1 + A_2 + A_3 - B_0) + 2\pi i A_1 \right]_{\alpha=n}.$$

It is now necessary to examine the quantity

$$(5.14) \quad W_n = \left[\frac{\partial}{\partial \alpha} (A_1 + A_2 + A_3 - B_0) \right]_{\alpha=n}.$$

Let $K = C(A_1) + C(A_2) + C(A_3)$, $\Gamma = K - C(B_0)$. Then

$$\begin{aligned} W_n &= \left[\frac{\partial}{\partial \alpha} \int_{\Gamma} t^{\alpha-2} \exp\left(\frac{1}{3} \lambda^{-2} t^3 + z t - \beta t^{-1}\right) dt \right]_{\alpha=n}, \\ &= \int_{\Gamma} t^{n-2} \log t \exp\left(\frac{1}{3} \lambda^{-2} t^3 + z t - \beta t^{-1}\right) dt. \end{aligned}$$

If we put

$$(5.15) \quad t = \lambda^{\frac{2}{3}} \sigma, \quad z = \lambda^{-\frac{2}{3}} \eta,$$

this becomes

$$\begin{aligned} W_n &= \lambda^{\frac{2}{3}(n-1)} \left[\frac{2}{3} (\log \lambda) \int_{\Gamma^*} \sigma^{n-2} \exp\left(\frac{\sigma^3}{3} + \eta \sigma - \frac{\beta}{\lambda^{\frac{2}{3}} \sigma}\right) d\sigma + \right. \\ &\quad \left. + \int_{\Gamma^*} \sigma^{n-2} \log \sigma \exp\left(\frac{\sigma^3}{3} + \eta \sigma - \frac{\beta}{\lambda^{\frac{2}{3}} \sigma}\right) d\sigma \right]. \end{aligned}$$

The first integral here vanishes because of CAUCHY'S theorem. Then

$$\begin{aligned} W_n &= \lambda^{\frac{2}{3}(n-1)} \left[\sum_{m=0}^{\infty} \frac{1}{m!} \left(\frac{-\beta}{\lambda^{\frac{2}{3}}}\right)^m \int_{K^*} \sigma^{n-2} \log \sigma \exp\left(\frac{\sigma^3}{3} + \eta \sigma\right) d\sigma - \right. \\ &\quad \left. - \int_{C^*(B_0)} \sigma^{n-2} \log \sigma \exp\left(\frac{\sigma^3}{3} + \eta \sigma - \frac{\beta}{\lambda^{\frac{2}{3}} \sigma}\right) d\sigma \right]. \end{aligned}$$

Reversing the transformation (5.15) in the last integral here, we find its contribution to be

$$\begin{aligned} &\frac{2}{3} (\log \lambda) \int_{C(B_0)} t^{n-2} \exp\left(\frac{1}{3} \lambda^{-2} t^3 + z t - \beta t^{-1}\right) dt - \\ &\quad - \int_{C(B_0)} t^{n-2} \log t \exp\left(\frac{1}{3} \lambda^{-2} t^3 + z t - \beta t^{-1}\right) dt \\ &= \frac{2}{3} (\log \lambda) B_0 - \left[\frac{\partial B_0}{\partial \alpha} \right]_{\alpha=n}. \end{aligned}$$

Finally, then,

$$(5.16) \quad \begin{aligned} W_n &= \lambda^{\frac{2}{3}(n-1)} \sum_{k=1}^3 \sum_{m=0}^{\infty} \frac{1}{m!} \left(\frac{-\beta}{\lambda^{\frac{2}{3}}}\right)^m \left[\frac{\partial}{\partial \alpha} g_k(\eta, \alpha - m) \right]_{\alpha=n} + \\ &\quad + \frac{2}{3} (\log \lambda) B_0 - \left[\frac{\partial B_0}{\partial \alpha} \right]_{\alpha=n}. \end{aligned}$$

The functions $g_k(\eta, \alpha)$ are defined by (4.24). The asymptotic expansions of the functions B_k when α is an integer n are then given by (5.16) and (5.13).

By taking $\alpha=0$, $\beta=1$, these results may be shown to agree with those of WASOW ([2], p. 227, equations (2.8a, b, c, d)). The relations connecting his

solutions U_1, U_2, U_3 with those considered here are

$$U_1 = B_3 - B_0,$$

$$U_2 = B_1,$$

$$U_3 = B_2.$$

6. A set of independent solutions

We consider here the set of solutions B_0, B_3, A_1, A_2 . Similar results will hold for any set B_0, B_i, A_j, A_k , where no two of i, j, k are equal. The solution B_0 is balanced everywhere in S . The solution B_3 is balanced in $S - S_3$ and is dominant in S_3 . The solution A_j is subdominant in S_j and is dominant in the other two sectors. This set will be shown to be linearly independent for $|\lambda|$ sufficiently large.

Let

$$C(\alpha) = i \sqrt{\pi} e^{-\frac{3\pi i}{2} \left(\frac{\alpha}{3} + \frac{1}{2} \right)}.$$

Then for z in $S - S_3$ and $|\xi| \geq |\xi_0|$ we have

$$(6.1) \quad B_0 = b_0(z, \alpha) + O(\lambda^{-2}),$$

$$(6.2) \quad B_3 = b_3(z, \alpha) [1 + O(\xi^{-2})],$$

$$(6.3) \quad A_1 = -C(\alpha) \lambda^{\alpha - \frac{3}{2}} z^{\frac{\alpha}{2} - \frac{5}{4}} e^{-\frac{2}{3} i \lambda z^{\frac{3}{2}}} [1 + O(\xi^{-1})],$$

$$(6.4) \quad A_2 = -i e^{\pi i \alpha} C(\alpha) \lambda^{\alpha - \frac{3}{2}} z^{\frac{\alpha}{2} - \frac{5}{4}} e^{\frac{2}{3} i \lambda z^{\frac{3}{2}}} [1 + O(\xi^{-1})],$$

where in (6.2), (6.3), (6.4)

$$-\frac{4\pi}{3} - \frac{2}{3} \arg \lambda < \arg z < -\frac{2}{3} \arg \lambda.$$

Let $W(f_1, \dots, f_n)$ be the Wronskian of f_1, \dots, f_n with respect to z . Using (2.3) to evaluate the derivatives of the above functions, by calculation we find

$$(6.5) \quad \begin{aligned} W(B_0, B_3, A_1, A_2) &= 2\pi i \lambda^{2\alpha+2} z^\alpha W(b_0, b_3) [1 + O(\xi^{-1})], \\ &= 4\pi^2 e^{-2\pi i \alpha} \beta^{\alpha-1} \lambda^{2\alpha+2} [1 + O(\xi^{-1})]. \end{aligned}$$

In the case where $|\xi| < |\xi_0|$ we have

$$(6.6) \quad B_0 = -2\pi i e^{-\pi i \alpha} \beta^{\alpha-1} \frac{1}{\Gamma(\alpha)} + O(\lambda^{-\frac{3}{2}}),$$

$$(6.7) \quad A_j = \lambda^{\frac{3}{2}(\alpha-1)} [g_j(\eta, \alpha) + O(\lambda^{-\frac{3}{2}})] \quad (j = 1, 2, 3).$$

Using (5.12c) we find

$$(6.8) \quad W(B_0, B_3, A_1, A_2) = \frac{e^{-2\pi i \alpha}}{1 - e^{-2\pi i \alpha}} W(B_0, A_1, A_2, A_3).$$

In the case where α is an integer n , we must take the limiting value of this expression as α tends to n . By use of (6.6) and (6.7), this becomes

$$(6.9) \quad \frac{-2\pi i \beta^{\alpha-1}}{\Gamma(\alpha)} \frac{e^{-3\pi i \alpha}}{1 - e^{-2\pi i \alpha}} \lambda^{2\alpha+2} W_\eta(g'_1, g'_2, g'_3) [1 + O(\lambda^{-\frac{3}{2}})],$$

where W_η denotes the Wronskian with respect to the variable η and the primes denote differentiation with respect to η . With the aid of (4.29) there follows

$$(6.10) \quad W_\eta(g'_1, g'_2, g'_3) = (1 - \alpha) W_\eta(g_1, g_2, g_3).$$

Since the Wronskian on the right is independent of η , it may be evaluated at $\eta = 0$. The quantities $g_k(0, \alpha)$ may be expressed in terms of gamma functions ([9], p. 225). In fact,

$$g_k(0, \alpha) = \int_{C^*(A_k)} \sigma^{\alpha-2} e^{\frac{\sigma^3}{3}} d\sigma = 2\pi i \frac{3^{\frac{\alpha-4}{3}} e^{\frac{2\pi i}{3}(\alpha-1)(k-2)}}{\Gamma\left(\frac{4-\alpha}{3}\right)}.$$

By the identity ([9], p. 225)

$$\Gamma\left(\frac{2-\alpha}{3}\right) \Gamma\left(\frac{3-\alpha}{3}\right) \Gamma\left(\frac{4-\alpha}{3}\right) = \frac{2\pi \Gamma(2-\alpha)}{3^{\frac{2}{3}-\alpha}}$$

and (4.25) there follows

$$(6.11) \quad W_1(g_1, g_2, g_3) = \frac{-4\pi^2}{\Gamma(2-\alpha)}.$$

From (6.9), (6.10), (6.11) and the familiar identity

$$\Gamma(2-\alpha) \Gamma(\alpha) = \frac{\pi(1-\alpha)}{\sin \pi \alpha}$$

we obtain finally

$$(6.12) \quad W(B_0, B_3, A_1, A_2) = 4\pi^2 e^{-2\pi i \alpha} \beta^{\alpha-1} \lambda^{2\alpha+2} [1 + O(\lambda^{-\frac{2}{3}})].$$

Then from (6.5) and (6.12) we see that if $|\lambda|$ is sufficiently large, the solutions B_0, B_3, A_1, A_2 are linearly independent.

A summary of the asymptotic properties of these functions is given below. The behavior of the other functions may be calculated from this information by using the connection formulas (2.6).

(i) The case $|\xi| \geq |\xi_0|$.

Let $\xi = \frac{2}{3}i\lambda z^{\frac{1}{3}}$, and in the formulae below let $\arg z$ lie in the appropriate part of the range

$$-2\pi - \frac{2}{3} \arg \lambda < \arg z < \frac{2\pi}{3} - \frac{2}{3} \arg \lambda.$$

$$(6.13) \quad B_0 = b_0(z, \alpha) + O(\lambda^{-2}) \quad \text{for } z \text{ in } S,$$

where

$$b_0(z, \alpha) = -2\pi i e^{-\pi i \alpha} (\beta^{-\frac{1}{2}} z^{\frac{1}{3}})^{1-\alpha} J_{\alpha-1}(2\beta^{\frac{1}{2}} z^{\frac{1}{3}}),$$

$$(6.14) \quad A_1 = -C(\alpha) \lambda^{\alpha-\frac{3}{2}} z^{\frac{\alpha}{2}-\frac{5}{4}} e^{-\xi} [1 + O(\xi^{-1})] \quad \text{for } z \text{ in } S-C_1,$$

$$(6.15) \quad A_2 = -i \lambda^{\alpha-\frac{3}{2}} e^{\pi i \alpha} C(\alpha) z^{\frac{\alpha}{2}-\frac{5}{4}} e^{\xi} [1 + O(\xi^{-1})] \quad \text{for } z \text{ in } S-C_2,$$

$$(6.16) \quad B_3 = \pi i (e^{2\pi i} \beta^{-\frac{1}{2}} z^{\frac{1}{3}})^{1-\alpha} H_{1-\alpha}^{(1)}(2\beta^{\frac{1}{2}} z^{\frac{1}{3}}) [1 + O(\xi^{-2})] \quad \text{for } z \text{ in } S-S_3,$$

$$(6.17) \quad B_3 = -C(\alpha) \lambda^{\alpha-\frac{3}{2}} z^{\frac{\alpha}{2}-\frac{5}{4}} e^{-\xi} [1 + O(\xi^{-1})] + \\ + \pi i e^{-2\pi i \alpha} (\beta^{-\frac{1}{2}} z^{\frac{1}{3}})^{1-\alpha} H_{1-\alpha}^{(1)}(2\beta^{\frac{1}{2}} z^{\frac{1}{3}}) [1 + O(\xi^{-2})] \quad \text{for } z \text{ in } S_3-C_1.$$

The expressions for A_1 and B_3 near C_1 and for A_2 near C_2 are more complicated but may be found easily from the relations

$$B_3 = B_0 + B_1 - A_2,$$

$$(6.18) \quad A_1 = B_0 + (1 - e^{-2\pi i \alpha}) B_1 - A_2 - e^{-2\pi i \alpha} A_3,$$

$$A_2 = B_0 + (1 - e^{-2\pi i \alpha}) B_2 - A_1 - A_3.$$

These follow from (2.6).

(ii) The case $|\xi| < |\xi_0|$ (for z in S).

$$(6.19) \quad B_0 = b_0(z, \alpha) + O(\lambda^{-2}) = -2\pi i e^{-\pi i \alpha} \beta^{\alpha-1} \left[\frac{1}{\Gamma(\alpha)} - \frac{\beta \eta}{\Gamma(\alpha+1)} \lambda^{-\frac{2}{3}} + \dots \right],$$

$$(6.20) \quad A_j = \lambda^{\frac{2}{3}(\alpha-1)} [g_j(\eta, \alpha) + O(\lambda^{-\frac{2}{3}})],$$

$$(6.21) \quad B_3 = \frac{e^{-\pi i \alpha}}{2i \sin \pi \alpha} \{ \lambda^{\frac{2}{3}(\alpha-1)} [(e^{2\pi i \alpha} g_1 + g_2 + g_3) + O(\lambda^{-\frac{2}{3}})] - b_0(z, \alpha) + O(\lambda^{-2}) \}$$

when α is not an integer.

$$(6.22) \quad B_3 = \frac{1}{2\pi i} \lambda^{\frac{2}{3}(n-1)} \left\{ \left[2\pi i g_1(\eta, \alpha) + \sum_{k=1}^3 \frac{\partial g_k(\eta, \alpha)}{\partial \alpha} \right]_{\alpha=n} + O(\lambda^{-\frac{2}{3}}) \right\} + \\ + \frac{1}{3\pi i} (\log \lambda) [b_0(z, n) + O(\lambda^{-2})] - \frac{1}{2\pi i} \left[\frac{\partial b_0(z, \alpha)}{\partial \alpha} \right]_{\alpha=n} + O(\lambda^{-2})$$

when α is an integer n .

In these formulae

$$g_k(\eta, \alpha) = \int_{C^*(A_k)} \sigma^{\alpha-2} \exp\left(\frac{\sigma^3}{3} + \eta \sigma\right) d\sigma,$$

$$g_k(0, \alpha) = 2\pi i \frac{3^{\frac{\alpha-4}{3}} e^{\frac{2\pi i}{3}(\alpha-1)(k-2)}}{\Gamma\left(\frac{4-\alpha}{3}\right)}.$$

For $z=0$, these formulae (ii) yield

$$(6.23) \quad B_0(0) \sim -2\pi i e^{-\pi i \alpha} \beta^{\alpha-1} \frac{1}{\Gamma(\alpha)},$$

$$(6.24) \quad A_k(0) \sim 2\pi i \frac{3^{\frac{\alpha-4}{3}} e^{\frac{2\pi i}{3}(\alpha-1)(k-2)}}{\Gamma\left(\frac{4-\alpha}{3}\right)} \lambda^{\frac{2}{3}(\alpha-1)},$$

$$(6.25) \quad B_3(0) \sim \pi \frac{3^{\frac{\alpha-4}{3}}}{\Gamma\left(\frac{4-\alpha}{3}\right)} \frac{e^{-\pi i \alpha} + 2 \cos \frac{\pi}{3}(\alpha+2)}{\sin \pi \alpha} \lambda^{\frac{2}{3}(\alpha-1)} + \\ + \pi \frac{e^{-2\pi i \alpha} \beta^{\alpha-1}}{\Gamma(\alpha) \sin \pi \alpha} \quad (\alpha \neq n),$$

$$(6.26) \quad B_3(0) \sim \lambda^{\frac{2}{3}(n-1)} \frac{3^{\frac{n-4}{3}}}{\Gamma\left(\frac{4-n}{3}\right)} \times \\ \times \left[2\pi i e^{\frac{2\pi i}{3}(1-n)} - \frac{4\pi}{3} \sin \frac{2\pi}{3}(n-1) + \frac{1}{3} \log 3 + \frac{1}{3} \psi\left(\frac{4-n}{3}\right) \right] - \\ - \frac{2}{3} e^{-n\pi i} \beta^{n-1} \frac{1}{\Gamma(n)} \log \lambda + \frac{d}{d\alpha} \left[e^{-\pi i \alpha} \beta^{\alpha-1} \frac{1}{\Gamma(\alpha)} \right]_{\alpha=n} \quad (\alpha = n).$$

(iii) *Bounds.*

Let ξ_j be the branch of $\xi = \frac{2}{3} i \lambda z^{\frac{2}{3}}$ which has a positive real part in S_j and a negative real part in the other two sectors. Let $E = E(z, \lambda)$ be a generic symbol for a uniformly bounded function of z and λ . Then the following relations hold.

For $|\xi| \geq |\xi_0|$ and $z \in S$,

$$(6.27a) \quad B_0 = E.$$

For $|\xi| < |\xi_0|$ and $z \in S$,

$$(6.27b) \quad B_0 = \begin{cases} E, & \alpha \neq 0, -1, -2, \dots \\ \lambda^{\frac{2}{3}(\alpha-1)} E, & \alpha = 0, -1, -2, \dots \end{cases}$$

For $|\xi| \geq |\xi_0|$,

$$(6.28a) \quad A_j e^{\operatorname{Re} \xi_j} = \lambda^{\frac{2}{3}(\alpha-1)} \xi^{\frac{\alpha}{3} - \frac{5}{6}} E, \quad z \in S - C_j \quad (j = 1, 2, 3).$$

For $z \in C_j$,

$$(6.28b) \quad A_j e^{\operatorname{Re} \xi_j} = \begin{cases} E, & \operatorname{Re} \alpha \leq 1, \alpha \neq 0, -1, -2, \dots \\ \lambda^{\frac{2}{3}(\alpha-1)} \xi^{-\frac{2}{3}(\alpha-1)} E, & 1 \leq \operatorname{Re} \alpha \leq \frac{3}{2}, \text{ or } \alpha = 0, -1, -2, \dots \\ \lambda^{\frac{2}{3}(\alpha-1)} \xi^{\frac{\alpha}{3} - \frac{5}{6}} E, & \operatorname{Re} \alpha \geq \frac{3}{2}. \end{cases}$$

For $|\xi| < |\xi_0|$ and $z \in S$,

$$(6.28c) \quad A_j e^{\operatorname{Re} \xi_j} = \lambda^{\frac{2}{3}(\alpha-1)} E.$$

For $|\xi| \geq |\xi_0|$ and $z \in S$ ($j = 1, 2, 3, j \neq k \neq l \neq j$),

$$(6.29a) \quad B_j e^{\frac{1}{2} \operatorname{Re} \{\xi_k + \xi_l\}} = \begin{cases} E, & \operatorname{Re} \alpha \leq 1, \alpha \neq 1 \\ \lambda^{\frac{2}{3}(\alpha-1)} \xi^{-\frac{2}{3}(\alpha-1)} E, & 1 < \operatorname{Re} \alpha < \frac{3}{2} \\ \lambda^{\frac{2}{3}(\alpha-1)} \xi^{\frac{\alpha}{3} - \frac{5}{6}} E, & \operatorname{Re} \alpha \geq \frac{3}{2} \\ \left(\log \frac{\lambda}{\xi} \right) E, & \alpha = 1. \end{cases}$$

For $|\xi| < |\xi_0|$ and $z \in S$,

$$(6.29b) \quad B_j e^{\frac{1}{2} \operatorname{Re} \{\xi_k + \xi_l\}} = \begin{cases} E, & \operatorname{Re} \alpha \leq 1, \alpha \neq 1 \\ (\log \lambda) E, & \alpha = 1 \\ \lambda^{\frac{2}{3}(\alpha-1)} E, & \operatorname{Re} \alpha > 1. \end{cases}$$

Acknowledgement. This work was sponsored in part by the Office of Naval Research. The author wishes to express his thanks to Professor C. C. LIN for suggesting this problem and for many helpful discussions with him during its development.

Note added in proof. Professor TRUESDELL has kindly pointed out the connection between equation (2.3) and his own work on the equation $\frac{\partial}{\partial z} F(z, \alpha) = F(z, \alpha + 1)$. (Annals of Mathematics Studies, Number 18, Princeton University Press, 1948.) Use of the theory for the F -equation would have shortened several of the formal steps in this paper.

References

- [1] WASOW, W.: A study of the solutions of the differential equation $y^{(4)} = \lambda^2(x y'' + y) = 0$ for large values of λ . *Ann. of Math.* **52**, 350—361 (1950).
- [2] WASOW, W.: Asymptotic solution of the differential equation of hydrodynamic stability in a domain containing a transition point. *Ann. of Math.* **58**, 222—252 (1953).
- [3] LANGER, R. E.: On the asymptotic solutions of a class of ordinary differential equations of the fourth order, with special reference to an equation of hydrodynamics. *Trans. Amer. Math. Soc.*, **84**, 144—191 (1957).
- [4] LANGER, R. E.: The asymptotic solutions of ordinary linear differential equations of the second order, with special reference to a turning point. *Trans. Amer. Math. Soc.* **67**, 461—490 (1949).
- [5] WASOW, W.: The complex asymptotic theory of a fourth order differential equation of hydrodynamics. *Ann. of Math.* **49**, 852—871 (1948).
- [6] WATSON, G. N.: *A Treatise on the Theory of Bessel Functions*, second edit. Cambridge 1944.
- [7] JEFFREYS, H., & B. S. JEFFREYS: *Methods of Mathematical Physics*, third edit. Cambridge 1956.
- [8] LANGER, R. E.: The solutions of the differential equation $v'''' + \lambda^2 z v' + 3\mu \lambda^2 v = 0$. *Duke Math. Journ.* **22**, 525—541 (1955).
- [9] COPSON, E. T.: *Theory of Functions of a Complex Variable*. Oxford 1935.

Massachusetts Institute of Technology
Cambridge, Massachusetts

(Received March 14, 1958)

Störungsrechnung bei Eigenwert- und Verzweigungsaufgaben

JOHANN SCHRÖDER

Vorgelegt von L. COLLATZ

Eine gegebene Eigenwert- oder Verzweigungsaufgabe kann man oft näherungsweise lösen, indem man von einem „benachbarten“ geschlossen lösbares Eigenwertproblem ausgeht. Um verschiedene Aufgabentypen (Matrizengleichungen, Differentialgleichungen, ...) gleichzeitig behandeln zu können, verwenden wir bei der Formulierung des allgemeinen Problems und in den Beweisen funktionalanalytische Begriffe und Methoden. Für bestimmte (in den §§ 8 und 9) als Beispiele behandelte Störungsprobleme bei Matrizen- und Differentialgleichungen werden die Ergebnisse dann jedoch so zusammengefaßt, daß man sie unmittelbar, also ohne funktionalanalytische Formulierung der Aufgabe, anwenden kann.

Wir behandeln Aufgaben der Form

$$A_0 \varphi - \lambda \varphi = B(\lambda, \varphi), \quad (0.1)$$

$$l \varphi = 1 \quad (0.2)$$

in einem Banachschen Raum. Gesucht sind also eine Zahl λ und ein Element φ , welche diesen Gleichungen genügen. A_0 sei ein linearer Operator, l ein beschränktes lineares Funktional. Die „benachbarte“ Aufgabe $A_0 \varphi - \lambda \varphi = 0$, $l \varphi = 1$ besitze einen einfachen Eigenwert $\lambda = \mu_0$ zum Eigenelement $\varphi = \psi_0$. Ist $B(\lambda, \varphi)$ linear in φ , so ist auch (0.1), (0.2) ein Eigenwertproblem, andernfalls liegt eine sogenannte Verzweigungsaufgabe vor.

Bei der üblichen Störungsrechnung führt man einen Störungsparameter ε ein, indem man B durch εB ersetzt, und berechnet λ und φ dann als Potenzreihen in ε

$$\lambda = \lambda(\varepsilon) = \mu_0 + \varepsilon \mu_1 + \varepsilon^2 \mu_2 + \dots, \quad \varphi = \varphi(\varepsilon) = \psi_0 + \varepsilon \psi_1 + \varepsilon^2 \psi_2 + \dots \quad (0.3)$$

Diese Methode wird für Eigenwertaufgaben z.B. in [2] und [3] beschrieben und in [6] auf Eigenwertaufgaben sowie in [7] (in etwas anderer Form) auch auf nicht-lineare Probleme angewendet. Wir benutzen hier stattdessen ein Iterationsverfahren. Das hat mancherlei Vorteile für die numerische Rechnung [z.B.: einfachere und bei jedem Schritt gleiche Rechenvorschrift, Selbstkorrektur des Verfahrens, weitgehend freie Wahl der Ausgangsnäherungen (siehe Beispiel 1), Konvergenz auch bei nichtkonvergenten Potenzreihen (0.3) (Beispiel 2)].

Dieser Aufsatz schließt an eine kürzlich erschienene Arbeit über das Newtonsche Verfahren [16] an, in welcher auch Iterationsverfahren der Form

$$u_{n+1} = u_n - [G'_0(x_n)]^{-1} G u_n \quad (n = 0, 1, 2, \dots) \quad (0.4)$$

zur Lösung einer Gleichung

$$G u = 0 \quad (0.5)$$

mit einem G „benachbarten“ Operator G_0 in einem abstrakten Raum \mathfrak{R} behandelt werden. Von diesem Raum \mathfrak{R} wird darin vorausgesetzt, daß jedem seiner Elemente u als verallgemeinerte Norm ein Element $v[u]$ eines halbgeordneten linearen Raumes \mathfrak{N} zugeordnet ist. Konvergenzaussagen und Fehlerabschätzungen für das Verfahren (0.4) werden dabei durch Vergleich mit einem Iterationsverfahren entsprechender Form

$$\varrho_{n+1} = \varrho_n - [\tilde{G}'_0(\varepsilon_n)]^{-1} \tilde{G} \varrho_n \quad (n = 0, 1, 2, \dots) \quad (0.6)$$

zur Lösung einer Gleichung

$$\tilde{G} \varrho = 0 \quad (0.7)$$

im Normenraum \mathfrak{N} erhalten.

Nach genauer Formulierung der Aufgabe (in § 1) definieren wir hier (in § 2) die Räume \mathfrak{R} und \mathfrak{N} , schreiben (in § 3) die Aufgabe (0.1), (0.2) in der Form (0.5) und leiten aus (0.4) das Iterationsverfahren her. Dann werden (in § 4) gewisse Majoranten ermittelt, welche in die anschließend (in § 5) untersuchte Vergleichsaufgabe (0.7) und das Vergleichsverfahren (0.6) eingehen. Um die Konvergenz des Verfahrens (0.6) zu beweisen, braucht man dabei nur zwei nichtlineare Ungleichungen für zwei reelle Unbekannte zu lösen. Das Ergebnis wird in § 6 zusammengefaßt. Besonders einfach wird der Rechengang, wenn A_0 ein selbstadjungierter Operator eines Hilbertschen Raumes ist.

Etwas ausführlicher behandeln wir (in § 7) noch den Spezialfall

$$(A_0 + A_1) \varphi = \lambda \varphi, \quad (0.8)$$

$$l \varphi = 1 \quad (0.9)$$

der Aufgabe (0.1), (0.2) in einem Hilbertschen Raum mit selbstadjungiertem Operator A_0 , linearem Operator A_1 und dem Funktional $l \varphi = (\varphi, \psi_0)$. Die Lösungen der genannten zwei Ungleichungen (bzw. Gleichungen) lassen sich hier (wie auch in anderen Fällen) in geschlossener Form angeben. Die Ergebnisse für diesen Fall enthalten u.a. auch die Fehlerabschätzungen für die Näherungen $\mu_0 + \mu_1$ und ψ_0 für λ bzw. φ , welche in [13] auf anderem Wege bewiesen wurden, und zwar einschließlich einer Verbesserung, auf deren Möglichkeit F. W. SCHÄFKE [12] hinwies (s. auch Beispiel 1).

Der Satz aus § 6 wird dann auf bestimmte Klassen von Eigenwertaufgaben bei Matrizen (in § 8) und gewöhnlichen Differentialgleichungen zweiter Ordnung (in § 9) angewendet. Unter anderem werden für Matrizeneigenwertaufgaben Formeln angegeben, mit deren Hilfe man den Einfluß von Abrundungsfehlern auf den Eigenwert abschätzen kann. — Die Genauigkeit der Abschätzungen prüfen wir (in § 10) numerisch an zwei einfachen Matrizeneigenwertaufgaben und erläutern am Beispiel eines Knickproblems (in § 11) ausführlich die Anwendung des Satzes aus § 6 auf nichtlineare Probleme.

Ein wesentliches Ziel dieser Arbeit ist es, für die betrachteten Probleme numerisch brauchbare Fehlerabschätzungen zu bekommen. Das Literaturverzeichnis nennt einige Arbeiten aus der Literatur über die Störungsrechnung, welche mit diesem Aufsatz in engerem Zusammenhang stehen und insbesondere auch Fehlerabschätzungen enthalten.

F. RELICH [8] untersuchte u. a. Störungsprobleme der Form

$$(A_0 + \varepsilon A_1 + \varepsilon^2 A_2 + \dots) \varphi = \lambda \varphi \quad (0.10)$$

in einem Hilbertschen Raum mit (linearen) Hermiteschen Operatoren A_i und selbstadjungiertem Operator A_0 und leitete dabei auch Fehlerabschätzungen für die Potenzreihen (0.3) her. B. v. SZ. NAGY [17] verbesserte diese Fehlerabschätzungen mit Hilfe einer andersartigen Beweismethode und erhielt entsprechende Ergebnisse auch für Probleme der Form (0.10) in einem Banachschen Raum [18]. Für Probleme der speziellen, aber praktisch wichtigsten Gestalt $(A_0 + \varepsilon A_1) \varphi = \lambda \varphi$ in einem Hilbertschen Raum mit selbstadjungiertem Operator A_0 , aber nicht notwendig Hermiteschem Operator A_1 wurden in [13] weiter verbesserte Fehlerabschätzungen für die Potenzreihen (0.3) angegeben. Von dem Operator A_1 wird darin verlangt, daß eine Ungleichung

$$\|A_1 \varphi\| \leq Q_1(\|\varphi\|, \|A_0 \varphi\|) \quad (0.11)$$

mit geeigneter Funktion $Q_1(s, t)$ gilt. Dies ist eine Verallgemeinerung der bei RELICH und SZ. NAGY benutzten Abschätzung $\|A_1 \varphi\| \leq a \|\varphi\| + b \|A_0 \varphi\|$.

P. ROSENBLOOM untersuchte in [9] Iterationsverfahren zur Lösung von Aufgaben der Gestalt (0.1), (0.2) in einem Banachschen Raum mit von λ unabhängiger Funktion B und anschließend zusammen mit H. BLOCK [1] auch solche Aufgaben, bei denen B von λ und φ abhängt¹. Gefordert wird dabei u. a., daß B auf einer bestimmten Teilmenge Lipschitzbedingungen

$$\|B(\lambda, \varphi) - B(\mu, \varphi)\| \leq D|\lambda - \mu|, \quad \|B(\lambda, \varphi) - B(\lambda, \psi)\| \leq E\|\varphi - \psi\| \quad (0.12)$$

genügt und l Eigenelement des adjungierten Operators A_0^* ist. Das behandelte Iterationsverfahren wird als Verfahren der Form

$$u_{n+1} = T u_n \quad (n = 0, 1, 2, \dots) \quad (0.13)$$

gedeutet und darauf der bekannte Fixpunktsatz für kontrahierende Abbildungen angewendet.

Unabhängig davon wurden in [15] Iterationsverfahren zur Lösung von Eigenwertaufgaben der Gestalt (0.8), (0.9) in einem Banachschen Raum untersucht, und zwar als Beispiel für die Anwendung allgemeinerer und genauerer Konvergenzkriterien für Iterationsverfahren der Art (0.13). Diese Untersuchungen wurden in [16] (siehe insbesondere § 6: Störungsrechnung) fortgesetzt, und die dort erhaltenen Ergebnisse werden hier nun angewendet. Dabei brauchen die Bedingungen (0.12), welche z. B. für die in § 11 behandelte Aufgabe nicht erfüllt sind, nicht gefordert zu werden, und das Funktional l ist weitgehend frei wählbar

¹ BLOCK und ROSENBLOOM untersuchten auch den Fall, daß A_0 nichtlinear ist. Sie benutzten andere Bezeichnungen, die von ihnen behandelten Probleme lassen sich jedoch in der Form (0.1), (0.2) schreiben.

(siehe ebenfalls § 11). $B(\lambda, \varphi)$ darf hier die Form einer Funktion $C(\lambda, \varphi, A_0\varphi, A_1\varphi, \dots, A_p\varphi)$ haben, welche in gewissem Sinne von einer reellwertigen Funktion \tilde{C} majorisiert wird. Die A_i bedeuten darin lineare Operatoren, für welche Ungleichungen der Art (0.11) gelten. Die hergeleiteten Fehlerabschätzungen sind einfacher und genauer als die in [1] angegebenen. Das hängt damit zusammen, daß der in (0.13) verwendete Operator T wegen des Gliedes $\lambda\varphi$ in (0.1) auch bei linearen Eigenwertaufgaben nichtlinear ist und eine Lipschitzbedingung für den Operator T immer nur in einem gewissen Teilbereich erfüllt ist.

F. W. SCHÄFKE behandelte in [12] (im Anschluß an frühere Arbeiten [10], [11] über speziellere Problemklassen) die Aufgabe (0.10) in einem Hilbertschen Raum mit selbstadjungiertem Operator A_0 und nicht notwendig Hermiteschen Operatoren A_1, A_2, \dots , für welche Abschätzungen der Art (0.11) gelten. Er leitete darin u.a. eine einfache, in gewissem Sinne nicht zu verbessernde Schranke für den Konvergenzradius der Reihen (0.3) her und verbesserte die in [13] bewiesenen Fehlerabschätzungen für $\mu_0 + \mu_1$ und ψ_0 (siehe dazu auch § 7). Im Beweis benutzte SCHÄFKE funktionentheoretische Überlegungen und den Fixpunktsatz für kontrahierende Abbildungen (in anderer Weise als BLOCK und ROSENBLUM).

Weitere Arbeiten über Störungsprobleme der Art (0.10) stammen z.B. von T. KATO [5] und F. WOLF [19].

§ 1. Aufgabenstellung

1.1. Es sei \mathfrak{S} ein reeller oder komplexer Banachscher Raum, dessen Elemente mit φ, ψ, \dots bezeichnet werden. $\|\varphi\|$ bedeute die Norm von φ . A_0 sei ein linearer Operator, welcher eine lineare Mannigfaltigkeit $\mathfrak{A} \subset \mathfrak{S}$ in \mathfrak{S} abbildet. Er besitze den Eigenwert μ_0 zum Eigelement ψ_0 :

$$A_0\psi_0 = \mu_0\psi_0.$$

Die weiteren Voraussetzungen für A_0 werden zunächst für den wichtigsten Spezialfall und erst anschließend für den allgemeinen Fall formuliert.

I. *Spezialfall*². \mathfrak{S} sei ein Hilbertscher Raum mit dem inneren Produkt (φ, ψ) und der Norm $\|\varphi\| = \sqrt{(\varphi, \varphi)}$. A_0 sei selbstadjungiert, μ_0 ein einfacher Eigenwert von A_0 und ψ_0 durch

$$(\psi_0, \psi_0) = 1$$

normiert. Das Spektrum von A_0 sei leer in einem Intervall

$$\mu_0 - d < \lambda < \mu_0 + d \quad \text{mit} \quad 0 < d < \infty$$

abgesehen von $\lambda = \mu_0$.

II. *Allgemeiner Fall*. ψ_0 sei so bestimmt, daß

$$l\psi_0 = 1 \tag{1.1}$$

gilt mit einem (der gegebenen Aufgabe entsprechend zu wählenden) auf \mathfrak{S} erklärten beschränkten linearen Funktional l .

² Dieser Fall wird im folgenden als „Spezialfall I“ zitiert.

Es gebe ein auf \mathfrak{S} definiertes beschränktes lineares Funktional Φ derart, daß die Gleichung

$$(A_0 - \mu_0 I) \varphi = \psi \quad (1.2)$$

(I = Einheitsoperator) für jedes $\psi \in \mathfrak{S}$ mit $\Phi \psi = 0$ eine Lösung φ besitzt, und die Gesamtheit der Lösungen dieser Gleichung habe die Gestalt

$$\varphi = c \psi_0 + R \psi$$

mit beliebiger Konstanten c und einem \mathfrak{S} in \mathfrak{U} abbildenden beschränkten linearen Operator R , für welchen

$$l R \psi = 0 \quad \text{bei} \quad \Phi \psi = 0$$

ist³.

Ferner gelte

$$\Phi \psi_0 \neq 0 \quad (1.3)$$

und

$$\Phi (A_0 - \mu_0 I) \varphi = 0 \quad \text{für} \quad \varphi \in \mathfrak{U}. \quad (1.4)$$

Im Spezialfall I setzen wir

$$l \varphi = \Phi \varphi = (\varphi, \psi_0) \quad (1.5)$$

und

$$R \psi = \int_{|\lambda - \mu_0| \geq d} \frac{1}{\lambda - \mu_0} dE_\lambda \quad (1.6)$$

mit der Spektralschar E_λ von A_0 . Diese Größen haben die im allgemeinen Fall geforderten Eigenschaften.

1.2. Gesucht sind Elemente $\varphi \in \mathfrak{U}$ und reelle bzw. komplexe Zahlen λ , welche die gestörte Gleichung

$$A_0 \varphi - \lambda \varphi = B(\lambda, \varphi) \quad (1.7)$$

erfüllen und der „Normierungsbedingung“

$$l \varphi = 1 \quad (1.8)$$

genügen. $B(\lambda, \varphi)$ habe die Form

$$B(\lambda, \varphi) = C(\lambda, \varphi, A_0 \varphi, A_1 \varphi, \dots, A_p \varphi) = C(\lambda, \varphi, A_i \varphi)^4. \quad (1.9)$$

Dabei bedeuten A_1, A_2, \dots, A_p lineare Operatoren mit dem Definitionsbereich \mathfrak{U} und Wertebereichen in \mathfrak{S} und C eine für reelle bzw. komplexe λ und $\varphi \in \mathfrak{U}$ definierte Funktion ihrer Variablen mit Werten aus \mathfrak{S} ⁵.

Die Operatoren A_i ($i=1, 2, \dots, p$) sollen vom Einheitsoperator I und A_0 in folgendem Sinne majorisiert werden. Es gebe für $s, t \geq 0$ erklärte stetige reell-

³ Das Funktional l ist weitgehend frei wählbar, R nicht eindeutig bestimmt. Φ und R kann man praktisch z. B. oft erhalten, indem man versucht, (1.2) bei zunächst beliebigem ψ zu lösen. Φ ist Eigenelement des adjungierten Operators A_0^* , R bei Differentialgleichungen z. B. ein Integraloperator mit einer verallgemeinerten Greenschen Funktion als Kern (siehe das Beispiel in § 11).

⁴ Gleichartige Variable werden in dieser Weise mit einem Index i zusammengefaßt, i durchlaufe dabei jeweils alle vorher noch nicht aufgeführten Werte. Zum Beispiel ist $C(\lambda, \varphi, A_0 \varphi, A_i \varphi) = C(\lambda, \varphi, A_i \varphi)$.

⁵ Der Fall, daß C nur auf einer Teilmenge dieses Definitionsbereiches erklärt ist, wird in Zusatz 2 in § 6 behandelt.

wertige Funktionen $Q_i(s, t)$ ($i = 1, 2, \dots, p$) mit den Eigenschaften:

$$\begin{aligned} 0 \leq Q_i(s, t) \leq Q_i(s', t') & \quad \text{für } 0 \leq s \leq s', 0 \leq t \leq t', \\ Q_i(\alpha s, \alpha t) = \alpha Q_i(s, t) & \quad \text{für } \alpha \geq 0; s, t \geq 0, \\ \|A_i \varphi\| \leq Q_i(\|\varphi\|, \|A_0 \varphi\|) & \quad \text{für } \varphi \in \mathfrak{A}. \end{aligned} \quad (1.10)$$

Ferner möge es eine für alle nichtnegativen Werte der reellen Variablen $\varrho^0, \varrho^1, \varrho^2, \dots, \varrho^{p+2}$ erklärte stetige reellwertige Funktion $\tilde{C}(\varrho^0, \varrho^1, \dots, \varrho^{p+2}) = \tilde{C}(\varrho^i)$ — eine Majorante für den Operator C — geben derart, daß gilt:

$$\begin{aligned} \tilde{C}(0, 0, \dots, 0) &= 0, \\ 0 \leq \tilde{C}(\varrho^i + \sigma^i) - \tilde{C}(\varrho^i) &\leq \tilde{C}(\xi^i + \eta^i) - \tilde{C}(\xi^i) \quad \text{für } 0 \leq \varrho^i \leq \xi^i, 0 \leq \sigma^i \leq \eta^i, \\ \|C(\lambda, \varphi, A_i \varphi) - C(\mu, \psi, A_i \psi)\| & \\ \leq \tilde{C}(|\lambda - \mu| + |\mu - \mu_0|, \|\varphi - \psi\| + \|\psi - \omega\|, \|A_i(\varphi - \psi)\| + \|A_i(\psi - \omega)\|) & \\ - \tilde{C}(|\mu - \mu_0|, \|\psi - \omega\|, \|A_i(\psi - \omega)\|) & \quad \text{für } \varphi, \psi \in \mathfrak{A} \end{aligned} \quad (1.11)$$

mit einem Bezugselement $\omega \in \mathfrak{A}$, für das im folgenden das Ausgangselement φ_0 des behandelten Iterationsverfahrens verwendet wird.

§ 2. Die Räume \mathfrak{N} und \mathfrak{R}

Es sei \mathfrak{R} die Gesamtheit der Kompositionen⁶ $u = \begin{pmatrix} \lambda_u \\ \varphi_u \end{pmatrix} = \begin{pmatrix} \lambda \\ \varphi \end{pmatrix}$ mit reellem bzw. komplexem λ und $\varphi \in \mathfrak{A}$ und \mathfrak{N} die Menge der reellen $(p+3)$ -dimensionalen Vektoren $\varrho = (\varrho^i)$ mit den Koordinaten ϱ^i ($i = 0, 1, 2, \dots, p+2$). Jedem $u \in \mathfrak{R}$ wird ein Vektor

$$v[u] = \begin{pmatrix} |\lambda| \\ \|\varphi\| \\ \|A_0 \varphi\| \\ \|A_1 \varphi\| \\ \vdots \\ \|A_p \varphi\| \end{pmatrix} = \begin{pmatrix} |\lambda| \\ \|\varphi\| \\ \|A_i \varphi\| \end{pmatrix} \in \mathfrak{N} \quad (p \geq 0)^7.$$

als (verallgemeinerte) Norm zugeordnet. Mit

$$\|\varrho\| = \max_i |\varrho^i|$$

$$\text{und} \quad \|u\| = \|v[u]\| \quad (2.1)$$

sind in \mathfrak{N} und \mathfrak{R} numerische Normen definiert.

\mathfrak{R} ist damit ein Banachscher Raum. Wir beweisen, daß auch der Raum \mathfrak{N} bezüglich der in ihm definierten numerischen Norm (2.1) vollständig ist.

Dazu werde zunächst gezeigt, daß der Operator A_0 abgeschlossen ist⁸. Es sei φ_n ($n = 1, 2, 3, \dots$) eine konvergente Folge aus \mathfrak{A} , für welche auch die Folge

⁶ Bei dem Element u lassen wir die Indizes an λ und φ häufig weg.

⁷ $v[u]$ habe also als Koordinaten mindestens $|\lambda|$, $\|\varphi\|$ und $\|A_0 \varphi\|$ und außerdem alle $\|A_i \varphi\|$ ($i = 1, 2, \dots$), welche in (1.9) wirklich vorkommen. Es können weitere $\|A_i \varphi\|$ hinzugenommen werden (wie z.B. in Nr. 9.2).

⁸ Im Spezialfall I braucht dies nicht bewiesen zu werden, da jeder selbstadjungierte Operator abgeschlossen ist.

$A_0 \varphi_n$ konvergiert. Wir setzen

$$\vartheta = \lim \varphi_n, \quad \vartheta = \lim \vartheta_n$$

mit

$$\vartheta_n = (A_0 - \mu_0 I) \varphi_n = (A_0 - \mu_0 I) (\varphi_n - l \varphi_n \cdot \psi_0).$$

Wegen (1.1) und (1.4) ist

$$l(\varphi_n - l \varphi_n \cdot \psi_0) = 0 \quad \text{und} \quad \Phi \vartheta_n = 0$$

und damit

$$\varphi_n - l \varphi_n \cdot \psi_0 = R \vartheta_n.$$

In der Grenze $n \rightarrow \infty$ erhält man

$$\varphi - l \varphi \cdot \psi_0 = R \vartheta \quad \text{und} \quad \Phi \vartheta = 0,$$

also ist $\varphi \in \mathfrak{A}$ und

$$(A_0 - \mu_0 I) \varphi = (A_0 - \mu_0 I) (\varphi - l \varphi \cdot \psi_0) = \vartheta,$$

so daß sich

$$\lim A_0 \varphi_n = \lim (A_0 - \mu_0 I) \varphi_n + \mu_0 \lim \varphi_n = \vartheta + \mu_0 \varphi = A_0 \varphi$$

ergibt. Der Operator A_0 ist also abgeschlossen.

Es sei nun $u_n (n=1, 2, 3, \dots)$ eine in sich konvergente Folge aus \mathfrak{R} , d.h. es gelte $\lim_{m, n \rightarrow \infty} \|u_m - u_n\| = 0$. Das bedeutet

$$|\lambda_m - \lambda_n| \rightarrow 0, \quad \|\varphi_m - \varphi_n\| \rightarrow 0, \quad \|A_i(\varphi_m - \varphi_n)\| \rightarrow 0 \quad (i=0, 1, 2, \dots, p; p \geq 0).$$

Daraus folgt wegen der Abgeschlossenheit von A_0 , daß die Folge φ_n gegen ein Element $\varphi \in \mathfrak{A}$ und die Folge $A_0 \varphi_n$ gegen $A_0 \varphi$ konvergiert. Aus (1.10) ergibt sich dann auch die Konvergenz der Folgen $A_i \varphi_n$ gegen $A_i \varphi (i=1, 2, \dots, p)$. Die Komposition u aus $\lambda = \lim \lambda_n$ und dem Element φ ist daher Grenzelement der Folge u_n im Sinne der Norm (2.1). Auch \mathfrak{R} ist also ein Banachscher Raum.

\mathfrak{R} und \mathfrak{S} haben damit die in [16], § 6 geforderten Eigenschaften.

§ 3. Herleiten des Iterationsverfahrens

3.1. Es seien G_0, G_1 und G die durch

$$G_0 u = A_0 \varphi - \lambda \varphi, \quad G_1 u = -B(\lambda, \varphi), \quad G = G_0 + G_1 \quad (3.1)$$

auf der Menge \mathfrak{D} der Elemente

$$u = \begin{pmatrix} \lambda \\ \varphi \end{pmatrix} \in \mathfrak{R} \quad \text{mit} \quad l \varphi = 1$$

definierten Operatoren. Die Aufgabe

$$G u = 0$$

ist dann dem Ausgangsproblem (1.7), (1.8) äquivalent.

Der Operator G_0 besitzt für jedes $u \in \mathfrak{D}$ die auf der abgeschlossenen Menge \mathfrak{L} der Elemente

$$h = \begin{pmatrix} \lambda_h \\ \varphi_h \end{pmatrix} \in \mathfrak{R} \quad \text{mit} \quad l \varphi_h = 0$$

definierte Ableitung $G'_{0(u)}h = A_0\varphi_h - \lambda_u\varphi_h - \lambda_h\varphi_u.$ (3.2)

Wir betrachten Iterationsverfahren der Form

$$G_0 u_n + G'_{0(x_n)}(u_{n+1} - u_n) + G_1 u_n = G u_n + G'_{0(x_n)}(u_{n+1} - u_n) = 0 \quad (n=0, 1, 2, \dots). \quad (3.3)$$

Der Unterschied zum gewöhnlichen Newtonschen Verfahren besteht darin, daß nur der Operator G_0 (bis zum linearen Glied) nach TAYLOR entwickelt wird, nicht der Gesamtoperator G , und außerdem die Ableitung an der Stelle x_n statt u_n genommen wird. Die x_n seien dabei aus der Menge \mathfrak{X} der Elemente

$$x = \begin{pmatrix} \mu_0 \\ \varphi_x \end{pmatrix} \quad \text{mit} \quad l\varphi_x = 1, \quad \Phi\varphi_x \neq 0,$$

und zwar wählen wir $\varphi_{x_n} = \varphi_{u_n}$ ⁹. Ausführlich lautet dann (3.3) mit $u_n = \begin{pmatrix} \lambda_n \\ \varphi_n \end{pmatrix}$:

$$\begin{aligned} (A_0 - \mu_0 I) \varphi_{n+1} &= (\lambda_{n+1} - \mu_0) \varphi_n + B(\lambda_n, \varphi_n), \\ l\varphi_{n+1} &= 1 \quad (n=0, 1, 2, \dots), \\ \varphi_0 &\in \mathfrak{A}, \quad l\varphi_0 = 1, \quad \Phi\varphi_0 \neq 0. \end{aligned} \quad (3.4)$$

Wendet man auf die erste dieser Gleichungen das Funktional Φ an, so ergibt sich die Formel für λ_{n+1} ⁹

$$\lambda_{n+1} = \mu_0 - \frac{1}{\Phi\varphi_n} \Phi B(\lambda_n, \varphi_n) \quad (n=0, 1, 2, \dots). \quad (3.5)$$

Nach dieser verhältnismäßig einfachen Vorschrift (3.4), (3.5) werden die Näherungen λ_n, φ_n praktisch berechnet. Im Falle $\Phi = l$ ist in (3.5) $\Phi\varphi_n = 1$ zu setzen

3.2. Für $x \in \mathfrak{X}$ besitzt die Gleichung

$$G'_{0(x)}h = (A_0 - \mu_0 I)\varphi_h - \lambda_h\varphi_x = \psi \quad (3.6)$$

bei beliebigem $\psi \in \mathfrak{S}$ eine eindeutige Lösung $h \in \mathfrak{L}$. Zunächst erhält man

$$\lambda_h = -\frac{1}{\Phi\varphi_x} \Phi\psi,$$

indem man Φ auf (3.6) anwendet. Wegen $\Phi(\psi + \lambda_h\varphi_x) = 0$ ergibt sich dann

$$\varphi_h = R\left(\psi - \frac{\Phi\psi}{\Phi\varphi_x} \cdot \varphi_x\right).$$

$[G'_{0(x)}]^{-1}$ ist also für $x \in \mathfrak{X}$ auf \mathfrak{S} durch

$$[G'_{0(x)}]^{-1}\psi = \begin{pmatrix} -\frac{\Phi\psi}{\Phi\varphi_x} \\ R\left(\psi - \frac{\Phi\psi}{\Phi\varphi_x} \cdot \varphi_x\right) \end{pmatrix}$$

definiert. Das Verfahren (3.3) läßt sich damit auch in der Form (0.4)

$$u_{n+1} = u_n - [G'_{0(x_n)}]^{-1} G u_n \quad (n=0, 1, 2, \dots) \quad (3.7)$$

schreiben.

⁹ Unter den Voraussetzungen des Satzes in § 6 liegen die Elemente $\varphi_{u_n} = \varphi_n$ in \mathfrak{X} , d.h. es gilt für sie auch $\Phi\varphi_n \neq 0$.

§ 4. Ermittlung von Majoranten

4.1. Bei der Untersuchung von Iterationsverfahren der Form (0.4) in [16] wird vorausgesetzt, daß die durch

$$H_0 = [G'_{0(w)}]^{-1} G_0, \quad H_1 = [G'_{0(w)}]^{-1} G_1$$

mit einem geeigneten Bezugselement w definierten Operatoren H_0 und H_1 in bestimmten Sinne durch im Normenraum \mathfrak{N} erklärte Operatoren \tilde{H}_0 und \tilde{H}_1 majorisiert werden. Solche Majoranten werden im folgenden für $w = x_0 = \begin{pmatrix} \mu_0 \\ \varphi_0 \end{pmatrix}$ ermittelt.

$\alpha_i (i = 1, 2, \dots, p+2)$ und $\beta_i (i = 0, 1, 2, 3)$ seien Konstanten, mit denen die folgenden Ungleichungen gelten:

$$\left. \begin{aligned} \|R\psi\| &\leq \alpha_1 \|\psi\|, & \alpha_1 > 0, \\ \|A_i R\psi\| &\leq \alpha_{i+2} \|\psi\| & (i = 0, 1, 2, \dots, p) \end{aligned} \right\} \quad \text{für } \psi \in \mathfrak{S} \quad \text{mit } \Phi\psi = 0,$$

$$\left. \begin{aligned} \left| \frac{\Phi\psi}{\Phi\varphi_0} \right| &\leq \beta_0 \|\psi\|, \\ \left\| \psi - \frac{\Phi\psi}{\Phi\varphi_0} \varphi_0 \right\| &\leq \beta_1 \|\psi\| \end{aligned} \right\} \quad \text{für } \psi \in \mathfrak{S} \quad \text{mit } l\psi = 0,$$

$$\left. \begin{aligned} \left| \frac{\Phi\psi}{\Phi\varphi_0} \right| &\leq \beta_2 \|\psi\|, \\ \left\| \psi - \frac{\Phi\psi}{\Phi\varphi_0} \varphi_0 \right\| &\leq \beta_3 \|\psi\| \end{aligned} \right\} \quad \text{für alle } \psi \in \mathfrak{S}. \quad (4.1)$$

Zum Beispiel kann man allgemein

$$\alpha_2 = 1 + |\mu_0| \alpha_1, \quad \alpha_{i+2} = Q_i(\alpha_1, \alpha_2) \quad (i = 1, 2, \dots, p)$$

verwenden, denn es ist für $\psi \in \mathfrak{S}$ mit $\Phi\psi = 0$

$$A_0 R\psi = (A_0 - \mu_0 I) R\psi + \mu_0 R\psi = \psi + \mu_0 R\psi,$$

also

$$\|A_0 R\psi\| \leq (1 + |\mu_0| \alpha_1) \|\psi\|,$$

und aus (1.10) folgt damit

$$\|A_i R\psi\| \leq Q_i(\|R\psi\|, \|A_0 R\psi\|) \leq Q_i(\alpha_1, \alpha_2) \|\psi\| \quad (i = 1, 2, \dots, p).$$

Im Falle $l = \Phi$ (also u.a. im Spezialfall I) kann man, wie unmittelbar zu sehen, außerdem

$$\beta_0 = 0, \quad \beta_1 = 1$$

setzen, und im Spezialfall I ferner

$$\alpha_1 = \frac{1}{d}, \quad \beta_2 = 1, \quad \beta_3 = 1 + \|\varphi_0 - \psi_0\|$$

benutzen. Denn im Spezialfall I gilt wegen (1.6)

$$\|R\| \leq \frac{1}{d},$$

und für $\psi \in \mathfrak{S}$ erhält man

$$\left| \frac{\Phi\psi}{\Phi\varphi_0} \right| = |(\psi, \psi_0)| \leq \|\psi\|,$$

sowie

$$\left\| \psi - \frac{\Phi \psi}{\Phi \varphi_0} \varphi_0 \right\| = \|\psi - (\psi, \varphi_0) \varphi_0\| \\ \leq \|\psi - (\psi, \varphi_0) \varphi_0\| + |(\psi, \varphi_0)| \|\varphi_0 - \varphi_0\| \leq (1 + \|\varphi_0 - \varphi_0\|) \|\psi\|.$$

Im Spezialfall I bereitet die Ermittlung der Konstanten α_i und β_i also keinerlei Schwierigkeiten.

(Die Forderungen (4.1) lassen sich abschwächen. Die Ungleichungen brauchen nur für Elemente ψ der Form

$$\psi = B(\lambda, \varphi) - B(\mu, \eta) \quad (\lambda, \mu \text{ reell bzw. komplex; } \varphi, \eta \in \mathfrak{A} \text{ mit } l\varphi = l\eta = 1) \quad (4.2)$$

zu gelten.)

Mit diesen Konstanten α_i, β_i schätzt man nun ab:

$$\nu[(G'_{0(w)})^{-1}\psi] \leq \|\psi\| \cdot v_0 \quad \text{für } \psi \in \mathfrak{S} \quad \text{mit } l\psi = 0, \\ \nu[(G'_{0(w)})^{-1}\psi] \leq \|\psi\| \cdot v_1 \quad \text{für } \psi \in \mathfrak{S},$$

worin v_0, v_1 die Vektoren

$$v_0 = \begin{pmatrix} \beta_0 \\ \alpha_1 \beta_1 \\ \alpha_2 \beta_1 \\ \dots \\ \alpha_{p+2} \beta_1 \end{pmatrix}, \quad v_1 = \begin{pmatrix} \beta_2 \\ \alpha_1 \beta_3 \\ \alpha_2 \beta_3 \\ \dots \\ \alpha_{p+2} \beta_3 \end{pmatrix}$$

bedeuten. Ferner ist nach (3.2) und (4.11) mit $\omega = \varphi_0$

$$\| [G'_{0(u)} - G'_{0(v)}] h \| \leq |\lambda_u - \lambda_v| \|\varphi_h\| + |\lambda_h| \|\varphi_u - \varphi_v\|, \\ \| G_1 u - G_1 v \| = \| B(\lambda, \varphi) - B(\mu, \psi) \| \\ \leq \tilde{C} (|\lambda - \mu| + |\mu - \mu_0|, \|\varphi - \psi\| + \|\psi - \varphi_0\|, \|A_i(\varphi - \psi)\| + \|A_i(\psi - \varphi_0)\|) - \\ - \tilde{C} (|\mu - \mu_0|, \|\psi - \varphi_0\|, \|A_i(\psi - \varphi_0)\|).$$

Insgesamt erhält man für $u, v \in \mathfrak{D}$ und $h \in \mathfrak{L}$

$$\nu[(H'_{0(u)} - H'_{0(v)}) h] \leq (\tilde{H}'_{0(v[u-v] + v[v-w])} - \tilde{H}'_{0(v[v-w])}) \nu[h], \\ \nu[H_1 u - H_1 v] \leq \tilde{H}_1(\nu[u-v] + \nu[v-w]) - \tilde{H}_1(\nu[v-w]) \quad (4.3)$$

mit den auf $\tilde{\mathfrak{D}} = \mathfrak{N}$ durch

$$\tilde{H}_0 \varrho = \varrho^0 \varrho^1 \cdot v_0, \quad \tilde{H}_1 \varrho = \tilde{C}(\varrho^i) \cdot v_1$$

erklärten Operatoren \tilde{H}_0 und \tilde{H}_1 .

4.2. Für Elemente $\xi \in \mathfrak{N}$ mit $\xi^0 = 0$ ist

$$[\tilde{H}'_{0(\xi)}]^j \sigma = \xi^1 \sigma^0 (\xi^1 \beta_0)^{j-1} \cdot v_0 \quad (j = 1, 2, 3, \dots).$$

In der Menge $\tilde{\mathfrak{Y}}$ der Elemente

$$\xi \in \mathfrak{N} \quad \text{mit} \quad \xi^0 = 0, \quad \xi^1 \geq 0 \quad \text{und} \quad \xi^1 \beta_0 < 1$$

existiert daher für $\sigma \in \mathfrak{N}$

$$[I - \tilde{H}'_{0(\xi)}]^{-1} \sigma = \sum_{j=1}^{\infty} [\tilde{H}'_{0(\xi)}]^{j-1} \sigma = \sigma + \frac{1}{1 - \xi^1 \beta_0} \xi^1 \sigma^0 \cdot v_0.$$

Die Operatoren \tilde{H}_0 und \tilde{H}_1 besitzen die in [16] (Nr. 6.1, 3.1) geforderten Eigenschaften, und alle dort vorkommenden Begriffe sind erklärt.

§ 5. Vergleichsaufgabe und Vergleichsverfahren

5.1. Zur Konvergenzuntersuchung und Fehlerabschätzung haben wir nach [16] (Satz und Zusatz 4) eine Vergleichsaufgabe mit dem Vergleichsoperator

$$\tilde{G} = \tilde{G}_0 + \tilde{G}_1$$

zu lösen, wobei \tilde{G}_0 und \tilde{G}_1 die durch

$$\begin{aligned} \tilde{G}_0 \varrho &= [I - \tilde{H}'_{0(\xi_0)}] \sigma_1 + \varrho_0 - \tilde{H}_0 \varrho_0 - \varrho + \tilde{H}_0 \varrho, \\ \tilde{G}_1 \varrho &= -\tilde{H}_1 \varrho_0 + \tilde{H}_1 \varrho \end{aligned}$$

auf \mathfrak{N} definierten Operatoren bedeuten und hier

$$\xi_0 = v[x_0 - w] = 0, \quad \varrho_0 = v[u_0 - w] = \begin{pmatrix} |\lambda_0 - \mu_0| \\ 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}$$

sowie

$$v[u_1 - u_0] = \begin{pmatrix} |\lambda_1 - \lambda_0| \\ \|\varphi_1 - \varphi_0\| \\ \|A_i(\varphi_1 - \varphi_0)\| \end{pmatrix} \leq \sigma_1 = (\sigma_1^k) \in \mathfrak{N}$$

ist. Gesucht wird ein Element $\tau \geq \varrho_0$ aus \mathfrak{N} mit

$$\tilde{G} \tau \leq 0, \quad (5.1)$$

also etwa eine Lösung der Gleichung $\tilde{G} \varrho = 0$.

(5.1) lautet ausführlich

$$\begin{aligned} \sigma_1^0 + \varrho_0^0 + \beta_0 \tau^0 \tau^1 + \beta_2 [\tilde{C}(\tau^i) - \tilde{C}(\varrho_0^i)] &\leq \tau^0, \\ \sigma_1^k + \alpha_k \beta_1 \tau^0 \tau^1 + \alpha_k \beta_3 [\tilde{C}(\tau^i) - \tilde{C}(\varrho_0^i)] &\leq \tau^k \quad (k = 1, 2, \dots, p+2). \end{aligned} \quad (5.2)$$

Sind τ^0 und τ^1 Zahlen, mit denen

$$\begin{aligned} \sigma_1^0 + \varrho_0^0 + \beta_0 \tau^0 \tau^1 + \beta_2 \left[\tilde{C}(\tau^0, \tau^1, \sigma_1^i + \frac{\alpha_i}{\alpha_1} (\tau^1 - \sigma_1^1)) - \tilde{C}(\varrho_0^0, 0, 0, \dots, 0) \right] &\leq \tau^0, \\ \sigma_1^1 + \alpha_1 \beta_1 \tau^0 \tau^1 + \alpha_1 \beta_3 \left[\tilde{C}(\tau^0, \tau^1, \sigma_1^i + \frac{\alpha_i}{\alpha_1} (\tau^1 - \sigma_1^1)) - \tilde{C}(\varrho_0^0, 0, 0, \dots, 0) \right] &\leq \tau^1 \end{aligned} \quad (5.3)$$

gilt, wobei i die Werte $2, 3, \dots, p+2$ durchläuft, so erfüllen die Zahlen

$$\tau^0, \tau^1, \tau^i = \sigma_1^i + \frac{\alpha_i}{\alpha_1} (\tau^1 - \sigma_1^1) \quad (i = 2, 3, \dots, p+2)$$

die Ungleichungen (5.2). Zur Ermittlung von τ genügt es also, die beiden nicht-linearen Ungleichungen (5.3) zu lösen.

5.2. Mit den Elementen

$$\xi_n = \begin{pmatrix} 0 \\ \varrho_n^1 \\ \varrho_n^2 \\ \dots \\ \varrho_n^{p+2} \end{pmatrix} \quad (n = 0, 1, 2, \dots)$$

lautet das Vergleichsverfahren (0.6)

$$\varrho_{n+1} = \varrho_n - [\tilde{G}'_{0(\xi_n)}]^{-1} \tilde{G} \varrho_n \quad (n = 0, 1, 2, \dots)$$

koordinatenweise geschrieben

$$\begin{aligned} \varrho_{n+1}^0 &= \frac{1}{1 - \beta_0 \varrho_n^1} \{ \sigma_1^0 + \varrho_0^0 + \beta_2 [\tilde{C}(\varrho_n^i) - \tilde{C}(\varrho_0^i)] \}, \\ \varrho_{n+1}^1 &= \sigma_1^1 + \alpha_1 \beta_1 \varrho_{n+1}^0 \varrho_n^1 + \alpha_1 \beta_3 [\tilde{C}(\varrho_n^i) - \tilde{C}(\varrho_0^i)], \\ \varrho_{n+1}^k &= \sigma_1^k + \frac{\alpha_k}{\alpha_1} (\varrho_{n+1}^1 - \sigma_1^1) \quad (k = 2, 3, \dots, p+2) \quad (n = 0, 1, 2, \dots). \end{aligned} \quad (5.4)$$

Soweit die ξ_n aus $\tilde{\mathfrak{Y}}$ sind, d.h. $\beta_0 \varrho_n^1 < 1$ gilt, genügen die Elemente u_n, x_n und ϱ_n, ξ_n den in [16] gestellten Forderungen

$$\begin{aligned} v[u_n - x_n] &\leq \varrho_n - \xi_n, \\ v[x_n - x_{n-1}] &\leq \xi_n - \xi_{n-1} \quad (n = 0, 1, 2, \dots; \quad x_{-1} = x_0, \quad \xi_{-1} = 0). \end{aligned} \quad (5.5)$$

Diese Ungleichungen folgen mit Hilfe vollständiger Induktion daraus, daß nach [16] (Nr. 2.4)

$$v[u_k - u_{k-1}] \leq \varrho_k - \varrho_{k-1} \quad (5.6)$$

gilt, soweit die vorkommenden Elemente mit (5.5) genügenden Elementen x_n und $\xi_n \in \tilde{\mathfrak{Y}}$ erklärt sind. Man vgl. auch das Beispiel in [16] (§ 4), wo ähnliche Elemente x_n und ξ_n verwendet werden.

Besitzen die Ungleichungen (5.3) eine Lösung τ^0, τ^1 mit $\beta_0 \tau^1 < 1, \tau^0 \geq \varrho_0^0, \tau_1^1 \geq \sigma_1^1$, so folgt aus dem Zusatz 4 in [16], daß die Iteration (5.4) unbeschränkt durchführbar ist und $\varrho_n \leq \tau$ (unter anderem also $\beta_0 \varrho_n^1 \leq \beta_0 \tau^1 < 1$) gilt. Da die Vektorfolge ϱ_n außerdem monoton zunimmt (vgl. (5.6)), konvergiert sie dann gegen einen Vektor $\varrho^* \leq \tau$. Dieser löst die Gleichung $\tilde{G}\varrho = 0$, was man einsieht, indem man in (5.4) den Grenzübergang $n \rightarrow \infty$ durchführt. Die dann entstehenden Gleichungen sind mit $\tilde{G}\varrho^* = 0$ identisch (vgl. (5.2)).

Existiert also eine solche Lösung der Ungleichungen (5.3), so sind alle Voraussetzungen des Satzes 4 in [16] erfüllt, und es ergibt sich das im nächsten Abschnitt formulierte Ergebnis.

5.3. Am genauesten werden die Fehlerschranken, wenn man $\sigma_1 = v[u_1 - u_0]$ setzt oder $v[u_1 - u_0]$ im Einzelfall möglichst gut abschätzt. Jedoch wollen wir diese Norm auch allgemein abschätzen, müssen dabei aber einen u.U. erheblichen Genauigkeitsverlust in Kauf nehmen. Eine solche Abschätzung für $v[u_1 - u_0]$ wird z.B. benötigt, wenn man φ_1 gar nicht berechnet hat.

Man erhält aus (3.4):

$$u_i - u_0 = \left(\frac{\mu_0 - \lambda_0 + \mu_1}{\psi_0 - \varphi_0 + R[\mu_1 \varphi_0 + B(\lambda_0, \varphi_0)]} \right) \quad \text{mit} \quad \mu_1 = - \frac{1}{\Phi \varphi_0} \Phi B(\lambda_0, \varphi_0),$$

also

$$v[u_1 - u_0] \leq \left(\frac{|\mu_0 - \lambda_0| + |\mu_1|}{\|\varphi_0 - \psi_0\| + \vartheta} \right) \quad \text{mit} \quad \vartheta = \alpha_1 \|\mu_1 \varphi_0 + B(\lambda_0, \varphi_0)\|. \quad (5.7)$$

Auf der rechten Seite dieser Ungleichung kommen nur Größen vor, welche man ohne Kenntnis von φ_1 , also ohne eine Gleichung lösen zu müssen, berechnen kann.

§ 6. Konvergenzaussagen und Fehlerabschätzungen (Ergebnis)

6.1. Wir fassen die aus den bisherigen Untersuchungen und [16] (Satz und Zusatz 4) folgenden Ergebnisse über das Iterationsverfahren (3.4), (3.5)

$$\lambda_{n+1} = \mu_0 - \frac{1}{\Phi \varphi_n} \Phi B(\lambda_n, \varphi_n),$$

$$(A_0 - \mu_0 I) \varphi_{n+1} = (\lambda_{n+1} - \mu_0) \varphi_n + B(\lambda_n, \varphi_n), \quad l \varphi_{n+1} = 1, \quad (6.1)$$

$$\varphi_0 \in \mathfrak{A} \quad \text{mit} \quad l \varphi_0 = 1, \quad \Phi \varphi_0 \neq 0 \quad (n = 0, 1, 2, \dots)$$

($l \varphi = \Phi \varphi = (\varphi, \psi_0)$ im Spezialfall I) zur Lösung der Aufgabe (1.7), (1.8)

$$A_0 \varphi - \lambda \varphi = B(\lambda, \varphi), \quad l \varphi = 1 \quad (6.2)$$

zusammen. Die in § 1 genannten Voraussetzungen seien dabei erfüllt.

Zur Fehlerabschätzung hat man folgende Rechnungen durchzuführen:

1. Man schätze

- a) die Operatoren A_i ($i = 1, 2, \dots, p$) in der Form (1.10) und
- b) den Operator C in der Form (1.11) mit $\omega = \varphi_0$ ab.

2. Man ermittle mit Hilfe der Tabellen 1 und 2 die Konstanten α_i ($i = 1, 2, \dots, p+2$), β_i ($i = 0, 1, 2, 3$) sowie die Schranken ϱ_0^0 und σ_1^i ($i = 0, 1, 2, \dots, p+2$).

3. Man stelle fest, ob die Ungleichungen (6.3) eine Lösung besitzen.

Satz. *Gibt es Zahlen τ^0, τ^1 mit $\tau^0 \geq \varrho_0^0$, $\tau^1 \geq \sigma_1^{10}$, $\beta_0 \tau^1 < 1$, welche die Ungleichungen¹¹*

$$\sigma_1^0 + \varrho_0^0 + \beta_0 \tau^0 \tau^1 + \beta_2 \left[\tilde{C} \left(\tau^0, \tau^1, \sigma_1^i + \frac{\alpha_i}{\alpha_1} (\tau^1 - \sigma_1^1) \right) - \tilde{C}(\varrho_0^0, 0, 0, \dots, 0) \right] \leq \tau^0,$$

$$\sigma_1^1 + \alpha_1 \beta_1 \tau^0 \tau^1 + \alpha_1 \beta_3 \left[\tilde{C} \left(\tau^0, \tau^1, \sigma_1^i + \frac{\alpha_i}{\alpha_1} (\tau^1 - \sigma_1^1) \right) - \tilde{C}(\varrho_0^0, 0, 0, \dots, 0) \right] \leq \tau^1 \quad (6.3)$$

erfüllen (wobei i die Werte $2, 3, \dots, p+2$ durchläuft), so konvergieren¹² die dann durch die Iterationsvorschrift definierten Folgen λ_n, φ_n gegen eine Lösung λ^, φ^* der*

¹⁰ Es genügt, statt $\tau^1 \geq \sigma_1^1$ die Ungleichungen $\tau^1 \geq \sigma_1^1 - \frac{\alpha_i}{\alpha_1} \sigma_1^i$ ($i = 1, 2, \dots, p+2$) nachzuprüfen. $\tau^1 \geq \sigma_1^1$ folgt dann aus (6.3).

¹¹ Wenn wir im folgenden von den „Gleichungen (6.3)“ sprechen, so sind damit die Beziehungen (6.3) mit \leq gemeint.

¹² Die Konvergenz von Elementenfolgen ist dabei im Sinne der jeweils verwendeten Norm zu verstehen, $\lim \varphi_n = \varphi^*$ bedeutet also $\lim \|\varphi^* - \varphi_n\| = 0$.

Tabelle 1 zur Berechnung der Konstanten α_i und β_i

Die Konstanten sind — sofern für sie nicht explizite Formeln angegeben sind — so zu bestimmen¹³, daß die genannten Ungleichungen mit ihnen erfüllt sind. Besonders einfach wird die Bestimmung der Konstanten also im Spezialfall I (\mathfrak{S} ist ein Hilbertscher Raum, A_0 selbstadjungiert). Auch im Spezialfall I kann man natürlich die für den allgemeinen Fall genannten Ungleichungen zur Bestimmung der Konstanten verwenden und dadurch u. U. kleinere Werte erhalten, jedoch wird man dies wegen der damit verbundenen größeren Rechenarbeit gewöhnlich nicht tun.

	Spezialfall I	Fall $l = \Phi$	Allgemeiner Fall
α_1	$= \frac{1}{d}$		$\ R\psi\ \leq \alpha_1 \ \psi\ $ für $\psi \in \mathfrak{S}$ mit $\Phi\psi = 0$
α_2	$= 1 + \mu_0 /d$		$= 1 + \mu_0 $ α_1 (genauer: $\ A_0 R\psi\ \leq \alpha_2 \ \psi\ $ für $\psi \in \mathfrak{S}$ mit $\Phi\psi = 0$)
$\alpha_{i-2} (i = 1, 2, \dots, p)$	$= Q_i \left(\frac{1}{d}, 1 + \frac{ \mu_0 }{d} \right)$		$= Q_i(\alpha_1, \alpha_2)$ (genauer: $\ A_i R\psi\ \leq \alpha_{i+2} \ \psi\ $ für $\psi \in \mathfrak{S}$ mit $\Phi\psi = 0$)
β_0	$= 0$	$= 0$	$\left \frac{1}{\Phi\varphi_0} \Phi\psi \right \leq \beta_0 \ \psi\ $ für $\psi \in \mathfrak{S}$ mit $l\psi = 0$
β_1	$= 1$	$= 1$	$\left\ \psi - \frac{1}{\Phi\varphi_0} \Phi\psi \cdot \varphi_0 \right\ \leq \beta_1 \ \psi\ $ für $\psi \in \mathfrak{S}$ mit $l\psi = 0$
β_2	$= 1$		$\left \frac{1}{\Phi\varphi_0} \Phi\psi \right \leq \beta_2 \ \psi\ $ für $\psi \in \mathfrak{S}^{13}$
β_3	$= 1 + \ \varphi_0 - \psi_0\ $		$\left\ \psi - \frac{1}{\Phi\varphi_0} \Phi\psi \cdot \varphi_0 \right\ \leq \beta_3 \ \psi\ $ für $\psi \in \mathfrak{S}^{13}$

¹³ Die Ungleichungen für β_2 und β_3 brauchen nur für Elemente ψ der speziellen Form (4.2) zu gelten.

Tabelle 2 zur Ermittlung der Schranken ϱ_0^0 und σ_1^i

Die Schranken sind so zu bestimmen, daß die in Spalte 2 stehenden Ungleichungen erfüllt sind. Für den Fall, daß φ_1 nicht berechnet wurde, sind in den weiteren Spalten grobe Schranken in expliziter Form angegeben, welche φ_1 nicht enthalten.

Zu berechnende Schranken	Im allgemeinen zu verwendende Formeln	Grobe Schranken (zu verwenden, falls φ_1 nicht berechnet ist)	
		allgemeiner Fall	Spezialfall I mit $\lambda_0 = \mu_0, \quad \varphi_0 = \psi_0, \quad B(\lambda, \varphi) = -A_1 \varphi$
ϱ_0^0	$ \lambda_0 - \mu_0 \stackrel{(<)}{\leq} \varrho_0^0$	$= \lambda_0 - \mu_0 $	$= 0$
σ_1^0	$ \lambda_1 - \lambda_0 \stackrel{(<)}{\leq} \sigma_1^0$	$= \lambda_0 - \mu_0 + \mu_1 $	$= \mu_1 $
σ_1^1	$\ \varphi_1 - \varphi_0\ \leq \sigma_1^1$	$= \ \varphi_0 - \psi_0\ + \vartheta$	$= \vartheta$
σ_1^{i+2} ($i = 0, 1, 2, \dots, p$)	$\ A_i(\varphi_1 - \varphi_0)\ \leq \sigma_1^{i+2}$	$= \ A_i(\varphi_0 - \psi_0)\ + \frac{\alpha_{i+2}}{\alpha_1} \vartheta$	$ \left. \begin{aligned} \mu_1 &= (\mathcal{A}_1 \psi_0, \psi_0) \leq d \vartheta, \\ \vartheta &= \frac{1}{d} \sqrt{\ \mathcal{A}_1 \psi_0\ ^2 - \mu_1 ^2} \\ &\leq \ \mathcal{A}_1 \psi_0\ \leq \gamma \leq \alpha, \\ \gamma &= Q_1 \left(\frac{1}{d}, \frac{ \mu_0 }{d} \right), \\ \alpha &= Q_1 \left(\frac{1}{d}, 1 + \frac{ \mu_0 }{d} \right) \end{aligned} \right\} $

gegebenen Aufgabe, und es gelten die Fehlerabschätzungen

$$|\lambda^* - \lambda_1| \leq \tau^0 - \sigma_1^0 - \varrho_0^0, \quad \|\varphi^* - \varphi_1\| \leq \tau^1 - \sigma_1^1 \quad (6.4)$$

und allgemeiner

$$|\lambda^* - \lambda_n| \leq \tau^0 - \varrho_n^0, \quad \|\varphi^* - \varphi_n\| \leq \tau^1 - \varrho_n^1 \quad (n = 0, 1, 2, \dots) \quad (6.5)$$

mit den durch (5.4) definierten Zahlen ϱ_n^0, ϱ_n^1 .

Zusatz 1. Unter den Voraussetzungen des Satzes besitzen die Folgen $A_i \varphi_n$ die Grenzelemente $A_i \varphi^*$, und es gilt

$$\|A_i(\varphi^* - \varphi_n)\| \leq \frac{\alpha_{i+2}}{\alpha_1} (\tau^1 - \varrho_n^1) \quad (n = 0, 1, 2, \dots; i = 0, 1, 2, \dots, p). \quad (6.6)$$

Zusatz 2. Es sei \mathfrak{D}_λ ein offenes Intervall der λ -Achse bzw. ein offenes konvexes Gebiet der λ -Ebene, in welchem μ_0 liegt, und \mathfrak{D}_φ eine offene konvexe Teilmenge von Elementen $\varphi \in \mathfrak{A}$ mit $l\varphi = 1$, welche ψ_0 und φ_0 enthält. Ist $B(\lambda, \varphi) = C(\lambda, \varphi, A_i \varphi)$ nur für Zahlen $\lambda \in \mathfrak{D}_\lambda$ und Elemente $\varphi \in \mathfrak{D}_\varphi$ erklärt und die in (1.11) enthaltene Ungleichung für C in dieser Teilmenge erfüllt, so gelten die Aussagen des obigen Satzes, falls \mathfrak{D}_λ alle Zahlen λ mit $|\lambda - \lambda_1| \leq \tau^0 - \sigma_1^1 - \varrho_0^0$ und \mathfrak{D}_φ alle Elemente $\varphi \in \mathfrak{A}$ mit $l\varphi = 1$ und $\|\varphi - \varphi_1\| \leq \tau^1 - \sigma_1^1$, $\|A_i(\varphi - \varphi_1)\| \leq \frac{\alpha_{i+2}}{\alpha_1} (\tau^1 - \sigma_1^1)$ ($i = 0, 1, 2, \dots, p$) enthält.

Die Richtigkeit dieses Zusatzes 2 ergibt sich, wenn man in [16] als \mathfrak{D} die Menge der Elemente $u \in \mathfrak{H}$ mit $\lambda_u \in \mathfrak{D}_\lambda$, $\varphi_u \in \mathfrak{D}_\varphi$ und $l\varphi_u = 1$ verwendet.

Bemerkungen. Ist A_i beschränkt, so kann man $Q_i(s, t) = \|A_i\| s$ verwenden. Für nichtbeschränkte Störoperatoren bei gewöhnlichen Differentialgleichungen sind Abschätzungsfunktionen Q_i in [14] angegeben (s. auch § 9). — Majoranten \tilde{C} werden im folgenden für verschiedene Fälle hergeleitet. — Im Spezialfall I erfordert die Ermittlung der Konstanten α_i und β_i nur sehr geringe Rechenarbeit. — Für den Spezialfall I und ein Störungsglied der Form $B(\lambda, \varphi) = -A_1 \varphi$ werden die Gleichungen (6.3) in § 7 gelöst. Ist A_1 dabei hermetisch, so lassen sich die Fehlerschranken noch verbessern, wie in 7.3 gezeigt wird.

6.2. Es sei noch bemerkt, daß man die in dem Satz enthaltene Existenzaussage sowie die Fehlerabschätzung (6.4) für die erste Näherung auch auf einfacherem Wege bekommen kann. Das untersuchte Iterationsverfahren stimmt für $n = 0$ mit dem Verfahren

$$\lambda_{n+1} = \mu_0 - \frac{1}{\Phi \varphi_0} \Phi(\lambda_n - \mu_0) (\varphi_n - \varphi_0) - \frac{1}{\Phi \varphi_0} \Phi B(\lambda_n, \varphi_n),$$

$$(A_0 - \mu_0 I) \varphi_{n+1} = (\lambda_{n+1} - \mu_0) \varphi_0 + (\lambda_n - \mu_0) (\varphi_n - \varphi_0) + B(\lambda_n, \varphi_n), \quad l\varphi_{n+1} = 1 \quad (6.7)$$

$$\varphi_0 \in \mathfrak{A} \quad \text{mit} \quad l\varphi_0 = 1, \quad \Phi \varphi_0 \neq 0 \quad (n = 0, 1, 2, \dots)$$

überein, d.h. die jeweiligen ersten Näherungen sind gleich. Dieses letztere Verfahren kann man aber in der einfacheren Form

$$u_{n+1} = T u_n \quad (n = 0, 1, 2, \dots)$$

mit dem durch

$$Tu = \left(\begin{array}{c} \mu_0 - \frac{1}{\Phi \varphi_0} \Phi(\lambda - \mu_0) (\varphi - \varphi_0) - \frac{1}{\Phi \varphi_0} \Phi B(\lambda, \varphi) \\ \psi_0 + R \left[(\lambda - \mu_0) (\varphi - \varphi_0) - \frac{1}{\Phi \varphi_0} \Phi(\lambda - \mu_0) (\varphi - \varphi_0) \cdot \varphi_0 + \right. \\ \left. + B(\lambda, \varphi) - \frac{1}{\Phi \varphi_0} \Phi B(\lambda, \varphi) \cdot \varphi_0 \right] \end{array} \right)$$

auf \mathfrak{D} definierten Operator schreiben. Darauf läßt sich dann unmittelbar der Satz 1 von [16] anwenden, welcher solche Verfahren als Spezialfall mit behandelt.

Durch im wesentlichen die gleichen Abschätzungen wie im § 4 beweist man eine Ungleichung

$$\nu[Tu, Tv] \leq F[\nu[u-v], \nu[u-w], \nu[v-w]] \quad \text{mit} \quad w = \begin{pmatrix} \mu_0 \\ \varphi_0 \end{pmatrix} = x_0$$

der in [16], § 1 geforderten Form, und Satz und Zusatz 1 aus [16] ergeben dann, daß für das Verfahren (6.7) der oben formulierte Satz gilt, wenn die q_n jetzt die Näherungen des zu (6.7) gehörenden Vergleichsverfahrens

$$q_{n+1} = \sigma_1 + q_0 + q_n^0 q_n^1 v_0 + [\tilde{C}(q_n^i) - \tilde{C}(q_0^i)] v_1 \quad (n = 0, 1, 2, \dots)$$

bedeuten.

Andererseits bekommt man diese Ergebnisse für das (im allgemeinen langsamer konvergierende) Verfahren (6.7) auch durch geringe Variation unserer bisherigen Überlegungen. Man braucht nur $x_n = x_0$ und $\xi_n = 0$ ($n = 0, 1, 2, \dots$) zu setzen.

§ 7. Spezialfälle bei linearen Eigenwertaufgaben

Wir untersuchen einige Spezialfälle der allgemeinen Aufgabenstellung genauer.

7.1. Für

$$B(\lambda, \varphi) = -A_1 \varphi + \lambda A_2 \varphi$$

geht die Gleichung (1.7) über in

$$(A_0 + A_1) \varphi = \lambda (I + A_2) \varphi,$$

und man schätzt dafür ab:

$$\begin{aligned} \|B(\lambda, \varphi) - B(\mu, \psi)\| &\leq \|A_1 \varphi - A_1 \psi\| + |\mu_0| \|A_2 \varphi - A_2 \psi\| + |\lambda - \mu| \|A_2 \varphi_0\| + \\ &+ \frac{1}{2} [(|\lambda - \mu| + |\mu - \mu_0|) + |\mu - \mu_0|] \times \\ &\quad \times [(\|A_2 \varphi - A_2 \psi\| + \|A_2 \psi - A_2 \varphi_0\|) - \|A_2 \psi - A_2 \varphi_0\|] + \\ &+ \frac{1}{2} [(|\lambda - \mu| + |\mu - \mu_0|) - |\mu - \mu_0|] \times \\ &\quad \times [(\|A_2 \varphi - A_2 \psi\| + \|A_2 \psi - A_2 \varphi_0\|) + \|A_2 \psi - A_2 \varphi_0\|]. \end{aligned}$$

Also gilt (1.11) mit

$$\tilde{C}(q^i) = q^3 + |\mu_0| q^4 + q^0 \|A_2 \varphi_0\| + q^0 q^4. \quad (7.1)$$

Für $A_2 = 0$ erhält man die besonders einfache Majorante

$$\tilde{C}(q^i) = q^3. \quad (7.2)$$

7.2. Es seien nun die Voraussetzungen des Spezialfalles I erfüllt, und außerdem habe B die Form

$$B(\lambda, \varphi) = -A_1 \varphi.$$

Damit haben wir den in [13] behandelten Spezialfall vor uns. Hierfür lauten die Gleichungen (6.3)

$$\begin{aligned} a + \alpha [c - b + y] &= x, \\ b + x y + \alpha \beta [c - b + y] &= y, \end{aligned} \quad (7.3)$$

wenn

$$a = \frac{1}{d} (\sigma_1^0 + \varrho_0^0), \quad b = \sigma_1^1, \quad c = \frac{1}{\alpha d} \sigma_1^3, \quad \alpha = \alpha_3, \quad \beta = \beta_3 \quad (7.4)$$

und

$$x = \frac{1}{d} \tau^0, \quad y = \tau^1$$

gesetzt wird¹⁴. Man kann sie leicht lösen und erhält das

Ergebnis¹⁴. Gilt (im Spezialfall I bei $B(\lambda, \varphi) = -A_1 \varphi$)

$$1 - (\alpha \beta + \alpha c + \alpha b + a) \geq 0 \quad (7.5)$$

und

$$r \geq 0 \quad \text{mit} \quad r = [1 - (\alpha \beta + \alpha(c - b) + a)]^2 - 4\alpha(b + \alpha\beta(c - b)),$$

so konvergieren die dann durch (6.1) definierten Folgen λ_n, φ_n gegen eine Lösung λ^*, φ^* der gegebenen Aufgabe (6.2), und es gilt die Fehlerabschätzung

$$|\lambda^* - \lambda_1| \leq d(x - a), \quad (7.6)$$

$$\|\varphi^* - \varphi_1\| \leq y - b, \quad \|A_0(\varphi^* - \varphi_1)\| \leq (d + |\mu_0|)(y - b), \quad \|A_1(\varphi^* - \varphi_1)\| \leq d\alpha(y - b)$$

mit

$$x = a + \alpha(c - b) + \alpha y, \quad y = \frac{1 - (\alpha \beta + \alpha(c - b) + a) - \sqrt{r}}{2\alpha}.$$

Setzen wir nun speziell $\lambda_0 = \mu_0, \varphi_0 = \psi_0$, so ist λ_1 mit der durch Potenzreihenentwicklung berechneten ersten Näherung $\mu_0 + \mu_1$ für den Eigenwert identisch, so daß wir die entsprechenden Fehlerabschätzungen vergleichen können. Es wird nach (5.7)

$$\begin{aligned} |\lambda_1 - \lambda_0| &\leq |\mu_1| \leq \|A_1 \psi_0\| \leq Q_1(\|\psi_0\|, \|A_0 \psi_0\|) = Q_1(1, |\mu_0|) = \gamma d \leq \alpha d, \\ \left. \begin{aligned} \|\varphi_1 - \varphi_0\| &\leq \vartheta \\ \|A_1(\varphi_1 - \varphi_0)\| &\leq \alpha d \vartheta \end{aligned} \right\} \vartheta = \frac{1}{d} \sqrt{\|A_1 \psi_0\|^2 - |\mu_1|^2} \leq \frac{1}{d} \|A_1 \psi_0\| \leq \gamma \leq \alpha \end{aligned} \quad (7.7)$$

mit

$$\mu_1 = (A_1 \psi_0, \psi_0), \quad \gamma = Q_1\left(\frac{1}{d}, \frac{|\mu_0|}{d}\right), \quad \alpha = Q_1\left(\frac{1}{d}, 1 + \frac{|\mu_0|}{d}\right).$$

Verwendet man diese Zahl $\alpha, \beta = 1$ und als σ_1^0, σ_1^1 und σ_1^3 in (7.4) die Schranken $|\mu_1|, \frac{1}{d} \|A_1 \psi_0\|$ und $\alpha \|A_1 \psi_0\|$ oder irgendwelche der in (7.7) stehenden größeren Schranken sowie $\varrho_0^0 = 0$, so erhält man eine Fehlerabschätzung (7.6) für λ_1 , welche in [13] auf anderem Wege bewiesen wurde. Die entsprechende Fehlerabschätzung mit der genaueren Größe ϑ leitete F. W. SCHÄFKE in [12] auf wieder andere Weise her¹⁵.

¹⁴ Den trivialen Fall $\alpha = 0$ lassen wir dabei außer acht.

¹⁵ Diese Verbesserung läßt sich auch durch geringe Variation des Beweises in [13] erreichen.

7.3. Für den in 7.2 behandelten Fall lassen sich die Abschätzungen verbessern, wenn A_1 Hermitisch ist. Es ist dann

$$H_1 u - H_1 v = -[G'_{0(w)}]^{-1} (B(\lambda, \varphi) - B(\mu, \psi)) = \begin{pmatrix} (A_1(\varphi - \psi), \psi_0) \\ R[A_1(\varphi - \psi) - (A_1(\varphi - \psi), \psi_0) \varphi_0] \end{pmatrix}$$

und

$$|(A_1(\varphi - \psi), \psi_0)| = |(\varphi - \psi, (A_1 - \mu_1 I) \psi_0)| \leq \|\varphi - \psi\| \|(A_1 - \mu_1 I) \psi_0\| = d\vartheta \|\varphi - \psi\|$$

(ϑ wie oben), so daß die Ungleichung (4.3) jetzt mit

$$\tilde{H}_1 \varrho = \begin{pmatrix} \vartheta \varrho^1 \\ \alpha_i \beta \varrho^3 \end{pmatrix} \quad \left(\text{statt } \tilde{H}_1 \varrho = \begin{pmatrix} \varrho^3 \\ \alpha_i \beta \varrho^3 \end{pmatrix} \right)$$

gilt ($\tilde{C} = \varrho^3$ nach (7.2)). Statt der Gleichungen (7.3) erhält man

$$a + \vartheta y = x, \quad b + xy + \alpha\beta[c - b + y] = y$$

und damit das

Ergebnis (für den Fall $\vartheta \neq 0$). Ist (im Spezialfall I bei $B(\lambda, \varphi) = -A_1 \varphi$) A_1 Hermitesch, so gilt das in 7.2 formulierte Ergebnis auch, wenn man die Konvergenzbedingung (7.5) durch

$$1 - (\alpha\beta + 2\vartheta b + a) \geq 0 \quad \text{und} \quad r \geq 0 \quad \text{mit} \quad r = [1 - (\alpha\beta + a)]^2 - 4\vartheta[b + \alpha\beta(c - b)]$$

ersetzt und die Größen

$$x = a + \vartheta y, \quad y = \frac{1 - (\alpha\beta + a) - \sqrt{r}}{2\vartheta}$$

verwendet¹⁶.

Auch in diesem Ergebnis sind wieder die für Hermiteschen Operator A_1 in [13] bewiesene sowie eine von F. W. SCHÄPFKE [12] hergeleitete Fehlerabschätzung für λ_1 enthalten.

§ 8. Eigenwertaufgaben bei Matrizen

8.1. Zu lösen sei die Eigenwertaufgabe

$$(M_0 + M_1) \varphi = \lambda (N_0 + N_1) \varphi, \quad (8.1)$$

worin φ einen m -dimensionalen Vektor bezeichnet und M_0, M_1, N_0 und N_1 $m \times m$ -Matrizen bedeuten. Die Matrizen M_0 und N_0 seien Hermitesch ($M_0 = \bar{M}'_0, N_0 = \bar{N}'_0$) und N_0 sei außerdem positiv definit ($\bar{\varphi}' N_0 \varphi > 0$ für $\varphi \neq 0$).

Die „ungestörte Aufgabe“

$$M_0 \varphi = \lambda N_0 \varphi \quad (8.2)$$

besitze einen einfachen Eigenwert μ_0 . ψ_0 sei ein zugehöriger durch

$$\bar{\psi}'_0 N_0 \psi_0 = 1$$

normierter Eigenvektor. Gesucht sind Lösungen λ, φ der Aufgabe (8.1) mit

$$\bar{\psi}'_0 N_0 \varphi = 1. \quad (8.3)$$

¹⁶ Die sich hiermit ergebenden Fehlerabschätzungen sind im allgemeinen genauer (oder jedenfalls nicht ungenauer) als die aus dem Ergebnis in Nr. 7.2 folgenden. Sicher ist dies z. B. der Fall für $c - b \geq 0$. Oft wird $c = b = 0$ sein, vgl. (7.7).

Jeder weitere Eigenwert des Problems (8.2) habe von μ_0 mindestens den Abstand d . Mit \varkappa bezeichnen wir den kleinsten Eigenwert der Matrix N_0 , mit $|M_1|$ und $|N_1|$ nichtnegative Konstanten, für welche

$$(\overline{M_1 \varphi})' M_1 \varphi \leq |M_1|^2 \bar{\varphi}' \varphi, \quad (\overline{N_1 \varphi})' N_1 \varphi \leq |N_1|^2 \bar{\varphi}' \varphi \quad (8.4)$$

gilt. Sind q_{ik} die Elemente der Matrix M_1 , so kann man z.B.

$$|M_1| = \sqrt{\sum_{i,k} |q_{ik}|^2} \quad (8.5)$$

oder

$$|M_1| = \sqrt{Z \cdot S}, \quad Z = \text{Max}_i \sum_k |q_{ik}|, \quad S = \text{Max}_k \sum_i |q_{ik}|$$

verwenden (entsprechend bei N_1).

Wendet man auf diese Aufgabe den Satz aus § 6 an, so erhält man das am Schluß der Nr. 8.2 formulierte unmittelbar anwendbare Ergebnis. Ein darin enthaltenes spezielleres Ergebnis ist außerdem in Nr. 8.3 genannt.

8.2. Dieses Problem läßt sich als Aufgabe der Form (4.7), (4.8) deuten, bei welcher die Voraussetzungen des Spezialfalles I erfüllt sind, und wir können dann den Satz aus § 6 anwenden. Sei der Hilbertsche Raum der Vektoren φ mit dem inneren Produkt

$$(\varphi, \psi) = \bar{\psi}' N_0 \varphi$$

und der Norm

$$\|\varphi\| = \sqrt{\bar{\varphi}' N_0 \varphi}. \quad (8.6)$$

Die Aufgabe

$$A_0 \varphi - \lambda \varphi = B(\lambda, \varphi), \quad (\varphi, \psi_0) = 1$$

mit

$$B(\lambda, \varphi) = -A_1 \varphi + \lambda A_2 \varphi, \quad A_0 = N_0^{-1} M_0, \quad A_1 = N_0^{-1} M_1, \quad A_2 = N_0^{-1} N_1$$

ist dem zu lösenden Problem (8.1), (8.3) äquivalent. A_0 ist dabei selbstadjungiert. Das zugehörige Iterationsverfahren (6.1) läßt sich in der Form

$$\begin{aligned} \lambda_{n+1} &= \mu_0 + \bar{\psi}_0' M_1 \varphi_n - \lambda_n \bar{\psi}_0' N_1 \varphi_n, \\ (M_0 - \mu_0 N_0) \varphi_{n+1} &= (\lambda_{n+1} - \mu_0) N_0 \varphi_n - M_1 \varphi_n + \lambda_n N_1 \varphi_n, \\ \bar{\psi}_0' N_0 \varphi_0 &= 1, \quad \bar{\psi}_0' N_0 \varphi_{n+1} = 1 \quad (n = 0, 1, 2, \dots) \end{aligned} \quad (8.7)$$

schreiben, wobei $l\varphi = \Phi\varphi = (\varphi, \psi_0)$ gesetzt ist.

Wir schätzen $B(\lambda, \varphi)$ in der erforderlichen Weise ab. Man erhält

$$\begin{aligned} \|A_1 \varphi\|^2 &= (\overline{M_1 \varphi})' N_0^{-1} M_1 \varphi \leq \frac{1}{\varkappa} (\overline{M_1 \varphi})' M_1 \varphi \leq \frac{1}{\varkappa} |M_1|^2 \bar{\varphi}' \varphi \leq \frac{1}{\varkappa^2} |M_1|^2 \bar{\varphi}' N_0 \varphi, \\ \|A_2 \varphi\|^2 &= (\overline{N_1 \varphi})' N_0^{-1} N_1 \varphi \leq \frac{1}{\varkappa} (\overline{N_1 \varphi})' N_1 \varphi \leq \frac{1}{\varkappa} |N_1|^2 \bar{\varphi}' \varphi \leq \frac{1}{\varkappa^2} |N_1|^2 \bar{\varphi}' N_0 \varphi, \end{aligned} \quad (8.8)$$

also

$$\|A_1 \varphi\| \leq \frac{|M_1|}{\varkappa} \|\varphi\|, \quad \|A_2 \varphi\| \leq \frac{|N_1|}{\varkappa} \|\varphi\| \quad (8.9)$$

und damit die Zahlen

$$\alpha_1 = \frac{1}{d}, \quad \alpha_2 = 1 + \frac{|\mu_0|}{d}, \quad \alpha_3 = \frac{|M_1|}{\varkappa d}, \quad \alpha_4 = \frac{|N_1|}{\varkappa d}. \quad (8.10)$$

Als \tilde{C} läßt sich nach 7.1 die in (7.1) angegebene Funktion

$$\tilde{C}(\varrho^0, \varrho^1, \dots, \varrho^4) = \varrho^3 + |\mu_0| \varrho^4 + \varrho^0 \|A_2 \varphi_0\| + \varrho^0 \varrho^4 \quad (8.11)$$

oder wegen (8.8) auch

$$\tilde{C}(\varrho^0, \varrho^1, \dots, \varrho^4) = \varrho^3 + |\mu_0| \varrho^4 + \varrho^0 \frac{|N_1|}{\kappa} \|\varphi_0\| + \varrho^0 \varrho^4 \quad (\|\varphi_0\| = 1 \text{ für } \varphi_0 = \psi_0) \quad (8.12)$$

verwenden.

Ergebnis. Für das Iterationsverfahren (8.7) zur Lösung der Aufgabe (8.1), (8.3) gilt (unter den in 8.1 genannten Voraussetzungen) der Satz aus § 6 mit $p=2$, wenn man darin

als $\|\varphi\|$ die Norm (8.6),

als \tilde{C} die Funktion (8.12) (oder (8.11)),

als α_i die Zahlen (8.10),

als β_i die in Tabelle 1 für den Spezialfall I genannten Konstanten,

als ϱ_0^i, σ_1^i ($i=0, 1, 3, 4$) die in Tabelle 2, Spalte 2 erklären und unter Benutzung von (8.8) zu berechnenden Schranken verwendet.

Die Gleichungen (6.3) sind hier leicht zu lösen, sie führen auf quadratische Gleichungen für τ^0 und τ^1 .

8.3. Wir leiten aus diesem Ergebnis noch eine einfache Fehlerschranke für $|\lambda^* - \mu_0|$ her, welche nur die oben erklärten Konstanten $d, \kappa, |M_1|, |N_1|$ enthält. Sie kann etwa dazu dienen, den Einfluß von Abrundungsfehlern bei den Matrizen M_0 und N_0 auf den Eigenwert μ_0 der Aufgabe (8.2) abzuschätzen.

Es ist

$$|(B(\mu_0, \psi_0), \psi_0)| \leq \|B(\mu_0, \psi_0)\|, \quad \|B(\mu_0, \psi_0) - (B(\mu_0, \psi_0), \psi_0) \psi_0\| \leq \|B(\mu_0, \psi_0)\|$$

und

$$\|B(\mu_0, \psi_0)\| \leq \|A_1 \psi_0\| + |\mu_0| \|A_2 \psi_0\| \leq q \cdot d$$

mit

$$q = \frac{1}{\kappa d} (|M_1| + |\mu_0| |N_1|).$$

Aus (5.7) folgt damit für $\lambda_0 = \mu_0, \varphi_0 = \psi_0$

$$v[u_1 - u_0] \leq \sigma_1 \quad \text{mit} \quad \sigma_1 = q \cdot d \begin{pmatrix} 1 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix}.$$

Mit diesen Werten σ_1^i und der Funktion (8.12) gehen die Gleichungen (6.3) über in

$$(q + s x)(1 + y) = x, \quad (q + s x)(1 + y) = y - x y \quad \text{bei} \quad x = \frac{\tau^0}{d}, \quad y = \tau^1, \quad s = \frac{|N_1|}{\kappa}.$$

Gilt $s + 2\sqrt{q} \leq 1$, so existiert eine Lösung x, y mit

$$x = \frac{1-s}{2} - \sqrt{\frac{(1-s)^2}{4} - q} \leq \frac{2q}{1-s}. \quad (8.13)$$

Der Satz aus § 6 liefert dann das

Ergebnis. *Gilt*

$$d|N_1| + 2\sqrt{\kappa d(|M_1| + |\mu_0||N_1|)} \leq \kappa d, \quad (8.14)$$

so besitzt die Aufgabe (8.1) einen Eigenwert λ^* mit

$$|\lambda^* - \mu_0| \leq 2 \frac{|M_1| + |\mu_0||N_1|}{\kappa - |N_1|} \quad (\text{genauer: } |\lambda^* - \mu_0| \leq dx). \quad (8.15)$$

Ist N_0 die Einheitsmatrix und $N_1 = 0$, so gehen diese Formeln (8.14), (8.15) über in

$$4|M_1| \leq d \quad \text{bzw.} \quad |\lambda^* - \mu_0| \leq 2|M_1|$$

$$\left(\text{genauer: } |\lambda^* - \mu_0| \leq \frac{d}{2} \left(1 - \sqrt{1 - 4 \frac{|M_1|}{d}} \right) \leq |M_1| \left(1 + 4 \frac{|M_1|}{d} \right) \right).$$

Für die Eigenwertaufgabe

$$(M_0 + M_1)\varphi = \lambda\varphi$$

mit normalen Matrizen M_0 und M_1 bewiesen A. J. HOFFMANN u. H. W. WIELANDT [4], daß allgemein (also ohne Zusatzbedingung)

$$\sum_{i=1}^m |\lambda^{*i} - \mu_0^i|^2 \leq |||M_1|||^2 \quad (8.16)$$

gilt, wenn μ_0^i die Eigenwerte der Matrix M_0 und λ^{*i} diejenigen der Matrix $M_0 + M_1$ bedeuten und man diese Eigenwerte in geeigneter Weise numeriert. $|||M_1|||$ ist dabei die in (8.5) genannte Konstante $|M_1|$. Wie HOFFMANN und WIELANDT bemerken, braucht eine solche Ungleichung bei nichtnormaler Matrix M_1 nicht zu gelten. Aus (8.16) folgert man für den einzelnen Eigenwert

$$|\lambda^* - \mu_0| \leq |||M_1|||.$$

Zum Vergleich sei bemerkt, daß bei normaler Matrix M_1 mit den Eigenwerten κ_i dieser Matrix

$$|||M_1||| = \sqrt{\sum_{i=0}^m |\kappa_i|^2}$$

gilt und die kleinstmögliche Konstante $|M_1|$ in (8.4)

$$|M_1| = \text{Max}_i |\kappa_i|$$

ist.

§ 9. Gewöhnliche Differentialgleichungen

9.1. Als Beispiel behandeln wir die Eigenwertaufgabe

$$-\varphi'' + M[\varphi] = \lambda\varphi + P(\lambda)N[\varphi], \quad (9.1)$$

$$\alpha\varphi(0) - \beta\varphi'(0) = 0, \quad \gamma\varphi(1) + \delta\varphi'(1) = 0. \quad (9.2)$$

Dabei bedeuten M und N lineare Differentialausdrücke

$$M[\varphi] = p_2(x)\varphi'' + p_1(x)\varphi' + p_0(x)\varphi,$$

$$N[\varphi] = q_1(x)\varphi' + q_0(x)\varphi$$

mit stetigen Funktionen $p_i(x)$ und $q_i(x)$ und

$$p_2(x) < 1 \quad (0 \leq x \leq 1), \quad (9.3)$$

$P(\lambda)$ ein Polynom in λ und $\alpha, \beta, \gamma, \delta$ Konstanten¹⁷ mit

$$\alpha, \beta, \gamma, \delta \geq 0, \quad \alpha + \beta > 0, \quad \gamma + \delta > 0.$$

c_i und d_i seien Zahlen, für welche (im Intervall $[0, 1]$)

$$|p_i(x)| \leq c_i \quad (i = 0, 1, 2), \quad |q_i(x)| \leq d_i \quad (i = 0, 1)$$

gilt.

Die „ungestörte“ Aufgabe

$$-\varphi'' = \lambda \varphi, \quad \alpha \varphi(0) - \beta \varphi'(0) = 0, \quad \gamma \varphi(1) + \delta \varphi'(1) = 0 \quad (9.4)$$

besitze $\lambda = \mu_0$ als einfachen Eigenwert, $\varphi = \psi_0(x)$ sei eine zugehörige, durch

$$\int_0^1 |\psi_0(x)|^2 dx = 1$$

normierte Eigenfunktion. Jeder weitere Eigenwert der Aufgabe (9.4) habe von μ_0 mindestens den Abstand d .

Gesucht wird eine Lösung $\lambda, \varphi(x)$ der Aufgabe (9.1), (9.2) mit stetigem $\varphi''(x)$ und

$$\int_0^1 \varphi(x) \overline{\psi_0(x)} dx = 1. \quad (9.5)$$

Das Ergebnis für dieses Problem ist am Schluß der Nr. 9.2 in unmittelbar anwendbarer Gestalt formuliert.

9.2. Wir schreiben die Aufgabe als Störungsproblem im Hilbertschen Raum $\mathfrak{S} = \mathfrak{L}_2[0, 1]$ mit dem inneren Produkt

$$(\varphi, \psi) = \int_0^1 \varphi(x) \overline{\psi(x)} dx$$

und der Norm

$$\|\varphi\| = \sqrt{\int_0^1 |\varphi(x)|^2 dx}. \quad (9.6)$$

\mathfrak{A} bedeute die Menge der Funktionen $\varphi(x)$ mit stetiger zweiter Ableitung, welche die Randbedingungen (9.2) erfüllen. A_0, A_1 und A_2 seien die auf \mathfrak{A} durch

$$A_0 \varphi = -\varphi'', \quad A_1 \varphi = M[\varphi], \quad A_2 \varphi = N[\varphi]$$

definierten Operatoren. Mit

$$B(\lambda, \varphi) = -A_1 \varphi + P(\lambda) A_2 \varphi = C(\lambda, A_1 \varphi, A_2 \varphi) \quad (9.7)$$

läßt sich das gegebene Problem in der Form

$$A_0 \varphi - \lambda \varphi = B(\lambda, \varphi), \quad (\varphi, \psi_0) = 1 \quad (9.8)$$

¹⁷ In § 7 wurden die Buchstaben α, β, γ bereits in anderer Bedeutung verwendet. Eine Verwechslung ist kaum zu befürchten.

schreiben. Die Tatsache, daß A_0 nicht selbstadjungiert ist im Sinne der in der Funktionalanalysis üblichen Definition¹⁸, erfordert im folgenden zusätzliche Überlegungen. Ohne Rücksicht darauf soll zunächst B abgeschätzt werden.

Nach [14] ist

$$\|A_1 \varphi\| \leq Q_1(\|\varphi\|, \|A_0 \varphi\|), \quad \|A_2 \varphi\| \leq Q_2(\|\varphi\|, \|A_0 \varphi\|) \quad (9.9)$$

mit

$$Q_1(s, t) = c_2 t + c_1 \sqrt{s} t + c_0 s, \quad Q_2(s, t) = d_1 \sqrt{s} t + d_0 s. \quad (9.10)$$

Ferner gilt für die durch (9.7) definierte Funktion C eine Ungleichung der in (1.11) geforderten Art mit $\omega = \varphi_0$ und

$$\tilde{C}(\varrho^0, \varrho^1, \dots, \varrho^4) = \varrho^3 + P(\varrho^0 + |\mu_0|)(\varrho^4 + \|N \varphi_0\|) - P(|\mu_0|) \|N \varphi_0\|. \quad (9.11)$$

Diese Funktion \tilde{C} leitet man auf ähnliche Weise her wie die im Beispiel von § 11 benutzte.

Das (6.1) entsprechende Iterationsverfahren für die hier behandelte Aufgabe lautet mit $\lambda \varphi = \Phi \varphi = (\varphi, \varphi_0)$

$$\begin{aligned} \lambda_{n+1} &= \mu_0 + \int_0^1 (M[\varphi_n] - P(\lambda_n) N[\varphi_n]) \bar{\varphi}_0 dx, \\ -\varphi_{n+1}'' - \mu_0 \varphi_{n+1} &= (\lambda_{n+1} - \mu_0) \varphi_n - M[\varphi_n] + P(\lambda_n) N[\varphi_n], \\ \alpha \varphi_{n+1}(0) - \beta \varphi_{n+1}'(0) &= 0, \quad \gamma \varphi_{n+1}(1) + \delta \varphi_{n+1}'(1) = 0, \\ \int_0^1 \varphi_0 \bar{\varphi}_0 dx &= 1, \quad \int_0^1 \varphi_{n+1} \bar{\varphi}_0 dx = 1 \quad (n = 0, 1, 2, \dots) \end{aligned} \quad (9.12)$$

($\varphi_0(x)$ erfülle dabei die Randbedingungen (9.2), und $\varphi_0''(x)$ sei stetig).

Der Operator A_0 besitzt ein in \mathfrak{S} vollständiges Orthonormalsystem von Eigenfunktionen. Er läßt sich zu einem selbstadjungierten Operator \hat{A}_0 mit dem Definitionsbereich \mathfrak{U} durch Abschließen fortsetzen. \mathfrak{U} besteht aus allen Grenzelementen $\varphi = \lim \varphi_n$ von Folgen $\varphi_n \in \mathfrak{A}$, für welche auch die Folgen $A_0 \varphi_n$ konvergieren, und $\hat{A}_0 \varphi$ wird dann durch $\hat{A}_0 \varphi = \lim A_0 \varphi_n$ definiert. Wegen (9.9) kann man auch die Operatoren A_1 und A_2 zu auf \mathfrak{U} erklärten Operatoren \hat{A}_1, \hat{A}_2 fortsetzen, indem man $\hat{A}_i \varphi = \lim A_i \varphi_n$ ($i = 1, 2$) definiert. Da der Operator \hat{A}_0 die gleichen Eigenwerte wie A_0 besitzt und die Funktionen Q_i und \tilde{C} stetig sind, sind für das (9.8) entsprechende Problem mit den fortgesetzten Operatoren alle im Spezialfall I geforderten Voraussetzungen erfüllt. Es gelten dafür also die Aussagen des Satzes aus § 6, dessen Voraussetzungen im folgenden erfüllt seien.

Die Elemente $\varphi_n \in \mathfrak{U}$ des entsprechenden Iterationsverfahrens sind sicher sogar aus \mathfrak{A} , wenn man $\varphi_0 \in \mathfrak{A}$ wählt, vgl. (9.12). Wir zeigen, daß auch φ^* in \mathfrak{U} liegt.

Es sei μ eine reelle Zahl, welche kein Eigenwert der ungestörten Aufgabe ist. $G(x, \xi)$ bedeute die zum Differentialausdruck

$$-\varphi'' - \mu \varphi \quad (9.13)$$

¹⁸ A_0 ist nicht abgeschlossen.

und den Randbedingungen (9.2) gehörende Greensche Funktion. Dann ist $(A_0 - \mu I)^{-1} \varphi$ auf der Menge der stetigen Funktionen φ durch

$$(A_0 - \mu I)^{-1} \varphi = \int_0^1 G(x, \xi) \varphi(\xi) d\xi$$

definiert und $(\hat{A}_0 - \mu I)^{-1} \varphi$ in der gleichen Form für $\varphi \in \mathfrak{E}$ erklärt. Daraus folgt, daß jede Funktion aus \mathfrak{U} als dem Wertebereich des Operators $(\hat{A}_0 - \mu I)^{-1}$ stetig ist und eine stetige Ableitung besitzt. φ^* genügt der Gleichung (9.1), wenn man dort $-\varphi''$ durch $\hat{A}_0 \varphi$ ersetzt. Wegen (9.3) ergibt sich daher, daß $(\hat{A}_0 - \mu I) \varphi^*$ stetig, also aus dem Wertebereich des Operators $(A_0 - \mu I)^{-1}$ ist, so daß φ^* aus \mathfrak{U} sein muß.

Man kann auch den Betrag $|\varphi^*(x) - \varphi_n(x)|$ abschätzen. Wir setzen $A_3 = A_0 - \mu I$ (μ wie oben). Dieser Operator kommt in (9.7) (ebenso wie auch A_0) explizit nicht vor, jedoch können wir den Zusatz 1 aus § 6 ohne weiteres für $p=3$ anwenden, also auch den Fehler $\|A_3(\varphi^* - \varphi_n)\|$ abschätzen. Man hat

$$\|A_3 R\| = \max_{\lambda} \frac{|\lambda - \mu|}{|\lambda - \mu_0|} \leq 1 + \frac{|\mu_0 - \mu|}{d} = \alpha_5,$$

wobei R den Operator (1.6) bedeutet und λ die Eigenwerte von A_0 , abgesehen von $\lambda = \mu_0$, durchläuft. Damit erhält man unter den Voraussetzungen des Satzes wegen $\frac{\alpha_5}{\alpha_1} = d + |\mu_0 - \mu|$ die Fehlerabschätzung

$$\|A_3(\varphi^* - \varphi_n)\| \leq (d + |\mu_0 - \mu|) (\tau^1 - \varrho_n^1) \quad (n = 0, 1, 2, \dots).$$

Mit Hilfe der Schwarzschen Ungleichung folgt

$$|\varphi^*(x) - \varphi_n(x)| = \left| \int_0^1 G(x, \xi) A_3(\varphi^* - \varphi_n) d\xi \right| \leq \sqrt{\int_0^1 |G(x, \xi)|^2 d\xi} \|A_3(\varphi^* - \varphi_n)\|$$

($G(x, \xi)$ wie oben), also

$$|\varphi^*(x) - \varphi_n(x)| = \sqrt{\int_0^1 |G(x, \xi)|^2 d\xi} (d + |\mu_0 - \mu|) (\tau^1 - \varrho_n^1) \quad (9.14)$$

$$(n = 0, 1, 2, \dots; \varrho_1^1 = \sigma_1^1).$$

Verwendet man in (9.14) $\tau^1 = \varrho^{*1}$, so ergibt sich wegen $\varrho_n \rightarrow \varrho^*$, daß die Folge $\varphi_n(x)$ gleichmäßig gegen $\varphi^*(x)$ konvergiert.

Ergebnis. Für das Iterationsverfahren (9.12) zur Lösung der Aufgabe (9.1), (9.2), (9.5) gilt (unter den in 9.1 genannten Voraussetzungen) der Satz aus § 6 mit $p=2$, wenn man darin

als $\|\varphi\|$ die Norm (9.6),
als Q_1, Q_2 die durch (9.10) erklärten Funktionen,
als \tilde{C} die in (9.11) genannte Funktion,
als α_i, β_i die in Tabelle 1 für den Spezialfall I aufgeführten Konstanten,
als ϱ_0^1, σ_1^1 ($i=0, 1, 3, 4$) die in Tabelle 2, Spalte 2 bei $A_1 \varphi = -M \varphi$, $A_2 \varphi = N \varphi$ erklärten Schranken verwendet.

Unter den Voraussetzungen dieses Satzes konvergiert die Folge $\varphi_n(x)$ sogar gleichmäßig (im Intervall $[0, 1]$) gegen $\varphi^*(x)$, und es gilt die Fehlerabschätzung (9.14), worin μ eine von den Eigenwerten der ungestörten Aufgabe (9.4) verschiedene reelle Zahl bedeutet und $G(x, \xi)$ die zum Differentialausdruck (9.13) und den Randbedingungen (9.2) gehörige Greensche Funktion ist.

§ 10. Zwei einfache Beispiele

Die zwei sehr einfachen Beispiele dieses Abschnitts dienen vor allem dazu, einen (mehr theoretischen) Einblick in die Genauigkeit der Abschätzungen zu erhalten. Wir verwenden dabei die Bezeichnungen des § 8.

10.1. Beispiel 1. Wir rechnen zum Vergleich die schon in [13] behandelte Matrizenaufgabe

$$(M_0 + M_1) \varphi = \lambda \varphi, \quad \bar{\psi}'_0 \varphi = 1 \quad (10.1)$$

mit

$$M_0 = \begin{pmatrix} 10 & 0 \\ 0 & -10 \end{pmatrix}, \quad M_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

und zwar soll die

$$\mu_0 = 10, \quad \psi_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

benachbarte Lösung ermittelt werden.

Die Tabelle 3 enthält die mit der üblichen Störungsrechnung (Potenzreihenentwicklung) erhaltenen Näherungen $\sum_{\nu=0}^n \mu_\nu$ für den Eigenwert (Spalte 2), die exacten Fehlerbeträge (Spalte 3) sowie die mit den Ergebnissen von [13] berechneten Fehlerschranken (Spalte 4)¹⁹.

Tabelle 3. Ergebnisse der üblichen Störungsrechnung
(Potenzreihenentwicklung nach ε , vgl. [13])

n	Näherungen $\tilde{\lambda}_n = \sum_{\nu=0}^n \mu_\nu$	$ \lambda^* - \tilde{\lambda}_n =$ (exact)	$ \lambda^* - \tilde{\lambda}_n \leq$ (nach Störungstheorie [13])
1	10	0,050 125 6...	0,052 778 2
2	9,95	0,000 125 6...	0,002 778 2
3	9,95	0,000 125 6...	0,000 278 2
4	9,949 875	0,000 000 6...	0,000 028 2
5	9,949 875	0,000 000 6...	0,000 003 2

In Spalte 2 der Tabelle 4 stehen die Näherungen λ_n , welche man mit μ_0, ψ_0 beginnend nach der Iterationsvorschrift (8.7) erhält, Spalte 3 enthält die exacten Fehlerbeträge. Die Fehlerschranken in Spalte 4 wurden nach (7.6) mit $\alpha = \frac{|M_1|}{d} = \frac{1}{20}$ und $\beta = 1 + \|\varphi_0 - \psi_0\|$ ermittelt. Dabei wurden die jeweils abzuschätzenden Größen als erste Näherung λ_1, φ_1 aufgefaßt, als φ_0 also die vorletzte Näherung

¹⁹ Wegen $\mu_1 = 0$ ergibt sich hier bei Verwendung der Größe ϑ in (7.7) keine Verbesserung.

verwendet und $\lambda_0 = \mu_0 = 10$ gesetzt (λ_0 kommt hier in der Iterationsvorschrift nicht vor, ist bei der Abschätzung also frei wählbar).

Tabelle 4. *Ergebnisse bei Anwendung des Iterationsverfahrens*

n	Näherungen λ_n für λ^*	$ \lambda^* - \lambda_n =$ (exact, zugleich Schranken nach (7.6) mit $\beta = \varphi_0^2 $)	$ \lambda^* - \lambda_n \leq$ (nach (7.6) mit $\beta = 1 + \ \varphi_0 - \psi_0\ $)
1	10	0,050 125 629...	0,052 778 2
2	9,95	0,000 125 629...	0,000 132 3
3	9,949 875	0,000 000 629...	0,000 000 663
4	9,949 874 374...	0,000 000 003...	0,000 000 003 4
5	9,949 874 371...		

Das Iterationsverfahren liefert bessere Werte, und die zugehörigen Fehlerabschätzungen sind genauer. Daß die ersten der genannten Fehlerschranken übereinstimmen, hängt mit der Einfachheit der Aufgabenstellung zusammen.

Ändert man in der Aufgabe M_1 ab zu

$$M_1 = \varepsilon \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

so wird

$$\lambda^* = \sqrt{100 - \varepsilon^2}, \quad \varphi^* = \left(\varepsilon^{-1} (10 - \sqrt{100 - \varepsilon^2}) \right).$$

Für $\lambda_0 = \mu_0$, $\varphi_0 = \psi_0$, $\alpha = \frac{|\varepsilon|}{20}$ und $\beta = 1$ ist die Konvergenzbedingung (7.5) erfüllt, falls $|\varepsilon| \leq \frac{20}{3}$ gilt. Eine von F.W. SCHÄRFKE [12] hergeleitete Konvergenzbedingung für die Potenzreihenentwicklung (0.3) ergibt die Forderung $|\varepsilon| < 10$.

Durch genauere Ermittlung von $\beta = \beta_3$ läßt sich hier jedoch als Konvergenzbedingung für das Iterationsverfahren auch $|\varepsilon| \leq 10$ erreichen. Für Vektoren χ mit $(\chi, \psi_0) = 0$, also Vektoren der Gestalt $\chi = \begin{pmatrix} 0 \\ \chi^2 \end{pmatrix}$ ist $A_1 \varphi = \varepsilon \begin{pmatrix} -\chi^2 \\ 0 \end{pmatrix}$ und

$$\|A_1 \chi - (A_1 \chi, \psi_0) \varphi_0\| = \left\| \varepsilon \begin{pmatrix} 0 \\ \chi^2 \varphi_0^2 \end{pmatrix} \right\| = |\varphi_0^2| \|A_1 \chi\| \quad \text{bei} \quad \varphi_0 = \begin{pmatrix} 1 \\ \varphi_0^2 \end{pmatrix}.$$

Die zweite der Ungleichungen (4.1) gilt daher für Elemente ψ der speziellen Form (4.2) $\psi = B(\lambda, \varphi) - B(\mu, \eta) = -A_1 \chi (\chi = \varphi - \eta)$ für $\beta_3 = \beta = |\varphi_0^2|$, im Spezialfall $\varphi_0 = \psi_0$ also für $\beta = 0$. Mit $\beta = 0$ folgt aus (7.5) für $\lambda_0 = \mu_0$, $\varphi_0 = \psi_0$ die Konvergenzbedingung $|\varepsilon| \leq 10$.

Verwendet man $\beta = |\varphi_0^2|$ bei der Fehlerabschätzung (7.6) für die Aufgabe (10.1), so stimmen die Fehlerschranken mit den exacten Fehlerbeträgen (Tabelle 4, Spalte 3) überein. Hieraus geht offenbar hervor, daß in § 4 wie auch in der vorangehenden Arbeit [16] vorsichtig abgeschätzt wurde. Praktisch wird man natürlich mit $\beta = 1 + \|\varphi_0 - \psi_0\|$ arbeiten.

10.2. Beispiel 2 ist eine ähnliche Matrizenaufgabe

$$(M_0 + M_1) \varphi = \lambda \varphi, \quad \bar{\psi}'_0 \varphi = 1$$

$$\text{mit } M_0 = \begin{pmatrix} 10 & 0 \\ 0 & -10 \end{pmatrix}, \quad M_1 = \varepsilon \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{und} \quad \varepsilon \geq 0.$$

Den Werten $\mu_0=10$, $\varphi_0=\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ „benachbart“ ist die Lösung

$$\lambda^*(\varepsilon) = 10 + \varepsilon z^*, \quad \varphi^*(\varepsilon) = \begin{pmatrix} 1 \\ z^* \end{pmatrix}, \quad z^* = \varepsilon^{-1}(-10 + \sqrt{100 + \varepsilon^2}).$$

Für $\lambda_0=\mu_0$, $\varphi_0=\begin{pmatrix} 1 \\ z_0 \end{pmatrix}$ erhält man

$$\lambda_{n+1} = 10 + \varepsilon z_n, \quad \varphi_n = \begin{pmatrix} 1 \\ z_n \end{pmatrix} \quad (n = 0, 1, 2, \dots),$$

wobei

$$z_{n+1} = f(z_n) \quad (n = 0, 1, 2, \dots) \quad \text{mit} \quad f(z) = \frac{\varepsilon}{20}(1 - z^2)$$

ist. Gilt

$$|f'(z^*)| < 1 \quad (10.2)$$

für die Lösung z^* der Gleichung $z = f(z)$, so konvergiert die Folge z_n gegen z^* , wenn man mit einem z_0 aus einer genügend kleinen Umgebung von z^* beginnt. (10.2) bedeutet

$$\varepsilon < 10 \cdot \sqrt{3}, \quad (10.3)$$

Während also die Potenzreihen von $\lambda^*(\varepsilon)$ und $\varphi^*(\varepsilon)$ den Konvergenzradius $\Omega=10$ besitzen, konvergieren die Näherungsfolgen der Iteration bei geeigneter Ausgangsnäherung φ_0 noch unter der Bedingung (10.3).

Diese Konvergenzbedingung kann man auch mit dem Satz aus § 6 erhalten. Es handelt sich hier um eine Aufgabe der speziellen in 7.2 untersuchten Art. Ebenso wie im Beispiel 1 läßt sich $\beta = |\varphi_0^2|$ verwenden. Die Größen a, b, c, β in (7.4) unterscheiden sich beliebig wenig von $\frac{1}{a} |(A_1 \varphi^*, \varphi_0)|$, 0, 0 bzw. $|\varphi^{*2}| = z^*$, wenn man $\lambda_0 = \mu_0$ setzt und φ_0 so wählt, daß

$$\|\varphi^* - \varphi_0\| < \delta \quad (10.4)$$

mit genügend kleinem δ ist. Auf den nicht schwierigen Beweis sei hier verzichtet. Gilt also

$$1 - \alpha z^* > |(A_1 \varphi^*, \varphi_0)| \quad \left(\alpha = \frac{\varepsilon}{20} \right), \quad (10.5)$$

so sind bei genügend kleinem δ in (10.4) auch die Konvergenzbedingungen (7.5) erfüllt. (10.5) führt aber gerade auf (10.3).

§ 11. Beispiel einer Verzweigungsaufgabe

Die §§ 8, 9 und 10 enthalten Anwendungen des Satzes aus § 6 auf Eigenwertaufgaben in einem Hilbertschen Raum. Hier wird der Satz nun zur Lösung einer nichtlinearen Aufgabe in einem Banachschen Raum benutzt.

11.1. Wir betrachten einen Stab, der an einem Ende eingespannt und am anderen Ende frei beweglich ist und welcher an dem freien Ende eine Last P trägt (Fig. 1). Die Auslenkungsfunktion $\varphi(x)$ (x = Bogenlänge) genügt (unter geeigneten Voraussetzungen) der Differentialgleichung

$$-\varphi'' = \lambda \varphi \sqrt{1 - (\varphi')^2} \quad (11.1)$$

und den Randbedingungen

$$\varphi(0) = \varphi'(1) = 0, \quad (11.2)$$

wobei λ bis auf einen Faktor die Last P bedeutet.

Der Zusammenhang zwischen der Auslenkung $\delta = \varphi(1)$ und λ wird (qualitativ) durch Fig. 2 veranschaulicht (gestrichelte Kurven). Die entsprechenden



Fig. 1. Knicken eines Stabes



Fig. 2. Zusammenhang zwischen Ausbiegung und Last

Kurven für das linearisierte Problem mit der Differentialgleichung

$$-\varphi'' = \lambda \varphi \quad (11.3)$$

sind in Fig. 2 punktiert gezeichnet. Diese vereinfachte Aufgabe liefert noch die richtigen Verzweigungsstellen (Knicklasten).

Wir wollen hier die Aufgabe²⁰

$$-\varphi'' = \lambda \varphi - \frac{1}{2} \lambda \varphi (\varphi')^2, \quad (11.4)$$

$$\varphi(0) = \varphi'(1) = 0, \quad \varphi(1) = \delta > 0 \quad (11.5)$$

bei gegebenem $\delta > 0$ lösen. Die Differentialgleichung entsteht aus (11.1), wenn man von der Taylorentwicklung der Wurzel nach $(\varphi')^2$ die ersten beiden Glieder berücksichtigt. Gesucht sind also λ und $\varphi(x)$ bei gegebenem $\delta = \varphi(1)$. Dabei gehen wir von der entsprechenden Aufgabe mit der Differentialgleichung (11.3) statt (11.4) als ungestörtem Problem aus. Dieses besitzt

$$\lambda - \lambda_0 = \frac{\pi^2}{4}, \quad \varphi = \psi_0 = \delta \cdot \sin \frac{\pi}{2} x$$

als Lösung. Eine benachbarte Lösung der gegebenen Aufgabe (11.4), (11.5) soll berechnet werden.

11.2. Zunächst werde die gestellte Aufgabe funktionalanalytisch formuliert. Es sei \mathfrak{S} der Banachsche Raum der im Intervall $[0, 1]$ stetigen Funktionen $\varphi(x)$ mit der Norm²¹

$$\|\varphi\| = \frac{1}{\delta} \max_{0 \leq x \leq 1} |\varphi(x)| \quad (11.6)$$

²⁰ Auf derartige Aufgaben wendete Pöschl in [7] die Störungsrechnung in etwas anderer Weise an.

²¹ Man könnte ohne weiteres $\|\varphi\| = \max |\varphi(x)|$ benutzen. Mit der Norm (11.6) lassen sich manche der auftretenden Gleichungen nur etwas bequemer schreiben.

und \mathfrak{A} die Menge der Funktionen $\varphi \in \mathfrak{S}$ mit stetiger zweiter Ableitung, welche den Randbedingungen (11.2) genügen. Mit den auf \mathfrak{A} definierten Operatoren

$$A_0 \varphi = -\varphi'', \quad A_1 \varphi = \varphi'$$

und

$$B(\lambda, \varphi) = C(\lambda, \varphi, A_1 \varphi) = -\frac{1}{2} \lambda \varphi (A_1 \varphi)^2 \quad (11.7)$$

sowie dem auf \mathfrak{S} durch

$$l\varphi = \frac{1}{\delta} \varphi(1)$$

erklärten Funktional l läßt sich die Aufgabe (11.4), (11.5) in der Form (1.7), (1.8) schreiben.

Die Aufgabe (1.2)

$$-\varphi'' - \frac{\pi^2}{4} \varphi = \psi(x), \quad \varphi(0) = \varphi'(1) = 0$$

ist für alle stetigen Funktionen $\psi(x)$ lösbar, für welche

$$\Phi \psi = 0 \quad \text{mit} \quad \Phi \psi = \int_0^1 \sin \frac{\pi}{2} x \psi(x) dx$$

gilt, und die allgemeine Lösung hat die Form

$$\varphi(x) = c \psi_0(x) + R\psi$$

mit beliebiger Konstanten c und

$$R\psi = \int_0^1 K(x, \xi) \psi(\xi) d\xi \quad \text{bei} \quad K(x, \xi) = \begin{cases} \frac{2}{\pi} \sin \frac{\pi}{2} x \cos \frac{\pi}{2} \xi & \text{für } x \leq \xi \\ \frac{2}{\pi} \sin \frac{\pi}{2} \xi \cos \frac{\pi}{2} x & \text{für } \xi \leq x. \end{cases}$$

Es ist

$$l R\psi = \frac{1}{\delta} (R\psi)_{x=1} = 0 \quad \text{und} \quad \Phi \psi_0 = \delta \int_0^1 \sin^2 \frac{\pi}{2} x dx = \frac{\delta}{2} > 0.$$

Ferner berechnet man mit partieller Integration unter Benutzung der Randbedingungen

$$\Phi(A_0 - \mu_0 I) \varphi = \int_0^1 \sin \frac{\pi}{2} x \left(-\varphi'' - \frac{\pi^2}{4} \varphi \right) dx = 0.$$

Die in 1.1 für den allgemeinen Fall genannten Voraussetzungen über A_0 sind daher mit den oben definierten Größen l , Φ und R erfüllt.

Man schätzt ab:

$$\|A_1 \varphi\| = \frac{1}{\delta} \max_x |\varphi'(x)| = \frac{1}{\delta} \max_x \left| \int_0^x \varphi''(\xi) d\xi \right| \leq \frac{1}{\delta} \max_{\xi} |\varphi''(\xi)| = \|A_0 \varphi\| \quad \text{für } \varphi \in \mathfrak{A},$$

kann also in (1.10) $Q_1(s, t) = t$ verwenden. Jedoch werden wir diese Funktion bei der Abschätzung nicht benutzen. Durch Ausmultiplizieren der auftretenden

Produkte $C(\dots, \dots, \dots)$ stellt man fest, daß

$$\begin{aligned}
 & |C(\lambda, \varphi, A_1 \varphi) - C(\mu, \psi, A_1 \psi)| \\
 &= |C((\lambda - \mu) + (\mu - \mu_0) + \mu_0, (\varphi - \psi) + (\psi - \psi_0) + \psi_0, \\
 &\quad (A_1 \varphi - A_1 \psi) + (A_1 \psi - A_1 \psi_0) + A_1 \psi_0) - \\
 &\quad C((\mu - \mu_0) + \mu_0, (\psi - \psi_0) + \psi_0, (A_1 \psi - A_1 \psi_0) + A_1 \psi_0)| \\
 &\leq C(|\lambda - \mu| + |\mu - \mu_0| + |\mu_0|, |\varphi - \psi| + |\psi - \psi_0| + |\psi_0|, \\
 &\quad |A_1 \varphi - A_1 \psi| + |A_1 \psi - A_1 \psi_0| + |A_1 \psi_0|) - \\
 &\quad - C(|\mu - \mu_0| + |\mu_0|, |\psi - \psi_0| + |\psi_0|, |A_1 \psi - A_1 \psi_0| + |A_1 \psi_0|) \\
 &\leq C(|\lambda - \mu| + |\mu - \mu_0| + |\mu_0|, \delta[|\varphi - \psi| + |\psi - \psi_0| + |\psi_0|], \\
 &\quad \delta[|A_1 \varphi - A_1 \psi| + |A_1 \psi - A_1 \psi_0| + |A_1 \psi_0|]) - \\
 &\quad - C(|\mu - \mu_0| + |\mu_0|, \delta[|\psi - \psi_0| + |\psi_0|], \delta[|A_1 \psi - A_1 \psi_0| + |A_1 \psi_0|])
 \end{aligned}$$

gilt²². Die Beziehungen (1.14) sind für die Funktion (11.7) demnach bei $\omega = \psi_0$ erfüllt mit

$$\tilde{C} = \frac{1}{2\delta} (\varrho^0 + |\mu_0|) \delta (\varrho^1 + |\psi_0|) [\delta (\varrho^3 + |A_1 \psi_0|)]^2 - \frac{1}{2\delta} |\mu_0| (\delta |\psi_0|) [\delta |A_1 \psi_0|]^2,$$

d.h.

$$\tilde{C}(\varrho^0, \varrho^1, \varrho^2, \varrho^3) = \frac{\delta^2}{2} \left[\left(\varrho^0 + \frac{\pi^2}{4} \right) (\varrho^1 + 1) \left(\varrho^3 + \frac{\pi}{2} \right)^2 - \frac{\pi^4}{16} \right]. \quad (11.8)$$

11.3. Für $\lambda_0 = \mu_0$, $\varphi_0 = \psi_0$ ergeben sich λ_1 und φ_1 aus den (6.1) entsprechenden Gleichungen

$$\begin{aligned}
 \lambda_1 &= \mu_0 + \mu_1 \quad \text{mit} \quad \mu_1 = -\frac{2}{\delta} \int_0^1 \sin \frac{\pi}{2} x \left(-\frac{1}{2} \right) \frac{\pi^2}{4} \delta \sin \frac{\pi}{2} x \left[\left(\delta \sin \frac{\pi}{2} x \right)' \right]^2 dx, \\
 -\varphi_1'' - \frac{\pi^2}{4} \varphi_1 &= (\lambda_1 - \mu_0) \delta \sin \frac{\pi}{2} x - \frac{1}{2} \frac{\pi^2}{4} \delta \sin \frac{\pi}{2} x \left[\left(\delta \sin \frac{\pi}{2} x \right)' \right]^2 = -\frac{\pi^4}{128} \delta^3 \sin \frac{3\pi}{2} x,
 \end{aligned}$$

$$\varphi_1(0) = \varphi_1'(1) = 0, \quad \varphi_1(1) = \delta$$

zu

$$\lambda_1 = \frac{\pi^2}{4} + \frac{\pi^4}{128} \delta^2, \quad \varphi_1(x) = \delta \sin \frac{\pi}{2} x - \frac{\pi^2}{256} \delta^3 \left(\sin \frac{\pi}{2} x + \sin \frac{3\pi}{2} x \right).$$

11.4. Wir ermitteln die in der Fehlerabschätzung benötigten Konstanten. Es ist

$$\begin{aligned}
 \|R\psi\| &= \frac{1}{\delta} \text{Max}_x \left| \int_0^1 K(x, \xi) \psi(\xi) d\xi \right| \leq \text{Max}_x \int_0^1 K(x, \xi) d\xi \cdot \|\psi\| = \alpha_1 \|\psi\| \\
 &\quad \text{mit} \quad \alpha_1 = \frac{4}{\pi^2} (\sqrt{2} - 1),
 \end{aligned}$$

²² In ähnlicher Weise kann man oft eine Majorante \tilde{C} ermitteln, wenn C ein Polynom seiner Variablen ist.

$$\|A_1 R \psi\| = \frac{1}{\delta} \operatorname{Max}_x \left| \int_0^1 \frac{\partial}{\partial x} K(x, \xi) \psi(\xi) d\xi \right| \leq \operatorname{Max}_x \int_0^1 \left| \frac{\partial}{\partial x} K(x, \xi) \right| d\xi \cdot \|\psi\| = \alpha_3 \|\psi\|$$

$$\text{mit } \alpha_3 = \frac{2}{\pi} (\sqrt{2} - 1),$$

$$\left| \frac{\Phi \psi}{\Phi \psi_0} \right| = \frac{2}{\delta} \left| \int_0^1 \sin \frac{\pi}{2} x \psi(x) dx \right| \leq 2 \int_0^1 \sin \frac{\pi}{2} x dx \|\psi\| = \beta_0 \|\psi\| \quad \text{mit } \beta_0 = \beta_2 = \frac{4}{\pi},$$

$$\begin{aligned} \left\| \psi - \frac{1}{\Phi \psi_0} \Phi \psi \cdot \psi_0 \right\| &\leq \frac{1}{\delta} \operatorname{Max}_x \left(|\psi(x)| + \left| \frac{2}{\delta} \int_0^1 \sin \frac{\pi}{2} \xi \psi(\xi) d\xi \right| \delta \sin \frac{\pi}{2} x \right) \\ &\leq \|\psi\| + 2 \int_0^1 \sin \frac{\pi}{2} \xi d\xi \cdot \|\psi\| = \beta_1 \|\psi\| \quad \text{mit } \beta_1 = \beta_3 = 1 + \frac{4}{\pi}, \end{aligned}$$

$$|\lambda_1 - \mu_0| = \frac{\pi^4}{128} \delta^2 = \sigma_1^0,$$

$$\|\varphi_1 - \psi_0\| = \frac{\pi^2}{256} \delta^2 \operatorname{Max}_x \left| \sin \frac{\pi}{2} x + \sin \frac{3\pi}{2} x \right| = \frac{\pi^2}{288} \sqrt{3} \delta^2 = \sigma_1^1,$$

$$\|A_1(\varphi_1 - \psi_0)\| = \frac{\pi^3}{512} \delta^2 \operatorname{Max}_x \left| \cos \frac{\pi}{2} x + 3 \cos \frac{3\pi}{2} x \right| = \frac{\pi^3}{288} \sqrt{2} \delta^2 = \sigma_1^3.$$

Zur Fehlerabschätzung sind die Gleichungen (6.3) mit diesen Konstanten $\alpha_i, \beta_i, \sigma_i^1$ und $\sigma_0^0 = 0$ sowie der Funktion \tilde{C} aus (11.8) zu lösen. Es wird $\tau^1 - \sigma_1^1 = \frac{\alpha_1 \beta_1}{\beta_0} (\tau^0 - \sigma_1^0)$. Man kann also τ^1 eliminieren und erhält dann leicht eine algebraische Gleichung 4. Grades für die Größe $a = \frac{1}{\delta^4} (\tau^0 - \sigma_1^0)$.

Damit ergibt sich nach dem Satz aus §6 die Fehlerabschätzung für $\delta = \frac{1}{10}$:

$$|\lambda^* - \lambda_1| = |\lambda^* - \mu_0 - \mu_1| = |\lambda^* - 2,46740 - 0,00761| = |\lambda^* - 2,47501| \leq 0,00024,$$

$$|\varphi^*(x) - \varphi_1(x)| = \left| \varphi^*(x) - 0,099961 \sin \frac{\pi}{2} x - 0,000039 \sin \frac{3\pi}{2} x \right| \leq 0,000008,$$

für $\delta = \frac{1}{5}$:

$$|\lambda^* - \lambda_1| = |\lambda^* - \mu_0 - \mu_1| = |\lambda^* - 2,4674 - 0,0304| = |\lambda^* - 2,4978| \leq 0,0047,$$

$$|\varphi^*(x) - \varphi_1(x)| = \left| \varphi^*(x) - 0,19969 \sin \frac{\pi}{2} x - 0,00031 \sin \frac{3\pi}{2} x \right| \leq 0,00028.$$

Eine formale Potenzreihenentwicklung von a nach δ liefert $a = 2,234 +$ Glieder mit höheren Potenzen, also

$$\begin{aligned} \tau^0 - \sigma_1^0 &= 2,234 \cdot \delta^4 \left\{ \right. \\ \tau^1 - \sigma_1^1 &= 0,670 \cdot \delta^4 \left. \right\} + \text{Glieder mit höheren Potenzen.} \end{aligned}$$

Die Potenzreihenentwicklung der Fehlerschranken für $|\lambda^* - \lambda_1|, \|\varphi^* - \varphi_1\|$ beginnt also mit der Potenz δ^4 , so daß — bei Konvergenz der Potenzreihen — das Verhältnis der Fehlerschranken zur Änderung $|\lambda_1 - \lambda_0|$ bzw. $\|\varphi_1 - \varphi_0\|$ mit δ wie δ^2 gegen Null strebt.

Literatur

- [1] BLOCK, H. D., & P. C. ROSENBLOOM: Perturbations of Non-Linear Eigenvalue Problems. Arch. Math. **7**, 172—183 (1956).
- [2] COLLATZ, L.: Eigenwertaufgaben mit technischen Anwendungen, S. 378ff. Leipzig 1949.
- [3] COURANT, R., u. D. HILBERT: Methoden der mathematischen Physik I, 2. Aufl., S. 296ff. Berlin 1931.
- [4] HOFFMANN, A. J., & H. W. WIELANDT: On the Variation of the Spectrum of a Normal Matrix. Nat. Bur. of Stand. Rep. 1578.
- [5] KATO, T.: On the convergence of the perturbation method I. Progress Theor. Physics **4**, 514—523 (1949); II. **5**, 95—101, 207—212 (1950).
- [6] MEYER ZUR CAPELLEN, W.: Methode zur angenäherten Lösung von Eigenwertproblemen mit Anwendungen auf Schwingungsprobleme. Ann. Phys. 5. Folge **8**, 297—352 (1931).
- [7] PÖSCHL, TH.: Über eine Methode zur angenäherten Lösung nichtlinearer Differentialgleichungen mit Anwendung auf die Berechnung der Durchbiegung bei der Knickung gerader Stäbe. Ing.-Arch. **9**, 34—41 (1938).
- [8] RELICH, F.: Störungstheorie der Spektralzerlegung. I. Math. Ann. **113**, 600—619 (1937); II. **113**, 677—685 (1937); III. **116**, 555—570 (1939); IV. **117**, 356—382 (1940/41); V. **118**, 462—484 (1942).
- [9] ROSENBLOOM, P. C.: Perturbation of Linear Operators in Banach Spaces. Arch. Math. **6**, 89—101 (1955).
- [10] SCHÄFKE, F. W.: Über Eigenwertprobleme mit zwei Parametern. Math. Nachr. **6**, 109—124 (1951).
- [11] SCHÄFKE, F. W.: Verbesserte Konvergenz- und Fehlerabschätzungen für die Störungsrechnung. Z. angew. Math. Mech. **33**, 255—259 (1953).
- [12] SCHÄFKE, F. W.: Zur Störungstheorie der Spektralzerlegung. Math. Ann. **133**, 219—234 (1957).
- [13] SCHRÖDER, J.: Fehlerabschätzungen zur Störungsrechnung bei linearen Eigenwertproblemen mit Operatoren eines Hilbertschen Raumes. Math. Nachr. **10**, 113—128 (1953).
- [14] SCHRÖDER, J.: Fehlerabschätzungen zur Störungsrechnung für lineare Eigenwertprobleme bei gewöhnlichen Differentialgleichungen. Z. angew. Math. Mech. **34**, 140—149 (1954).
- [15] SCHRÖDER, J.: Nichtlineare Majoranten beim Verfahren der schrittweisen Näherung. Arch. Math. **7**, 471—484 (1956).
- [16] SCHRÖDER, J.: Über das Newtonsche Verfahren. Arch. Rat. Mech. Anal. **1**, 154—180 (1957).
- [17] SZ. NAGY, B. v.: Perturbations des transformations autoadjointes dans l'espace de Hilbert. Comm. Math. Helvet. **19**, 347—366 (1946/47).
- [18] SZ. NAGY, B. v.: Perturbations des transformations linéaires fermées. Acta Sci. math. **14**, 125—137 (1952).
- [19] WOLF, F.: Analytic perturbation of operators in Banach spaces. Math. Ann. **124**, 317—333 (1952).

Mathematisches Institut
der Universität Hamburg

(Eingegangen am 31. März 1958)

*Note on a paper "Further Remarks on the Stress-Deformation Relations for Isotropic Materials"**

A. PIPKIN & R. S. RIVLIN

In § 11 of the above paper, the result is obtained that any polynomial scalar invariant under the group of orthogonal transformations of any number of 3×3 matrices is expressible as a polynomial in traces of products of the matrices. This result is in general valid only if all of the 3×3 matrices are symmetric. More generally, if the matrices may be non-symmetric, the argument given in the paper leads to the result that any polynomial scalar invariant under the group of orthogonal transformations is expressible as a polynomial in traces of products formed from the matrices and their transposes.

If one of the matrices is non-symmetric and the remaining matrices are symmetric and the polynomial scalar invariant is linear in the elements of the non-symmetric matrix, then it may be expressed as a polynomial in traces of products formed from the matrices. This is the theorem which is used in § 3 of a paper by A. E. GREEN & R. S. RIVLIN**.

The results presented in this note were obtained in the course of research sponsored by the Office of Ordnance Research, U. S. Army, under Contract No. DA-19-020-ORD-3487.

* J. Rational Mech. Anal. **4**, 681 (1955)

** Arch. Rational Mech. Anal. **1**, 1 (1957).

Brown University,
Providence, Rhode Island

(Received September 5, 1957)

Errata

Correction to A. E. GREEN & R. S. RIVLIN, *The Mechanics of Non-Linear Materials with Memory*, Vol. 1, 1—21 (1957)

Equation (3.12) does not follow strictly from the definition of F , since although b_{ij} is an arbitrary tensor, B_{ij} is not. We may, however, obtain equation (3.12) by the following argument.

Since F may be considered either as a linear function of B_{rs} or b_{ij} , we have, with (3.10),

$$\left(\frac{\partial x_i}{\partial X_r} \frac{\partial x_j}{\partial X_s} - \frac{1}{3} g_{rs} \delta_{ij} \right) f_{rs} = \frac{\partial F}{\partial b_{ij}} = \frac{\partial F}{\partial B_{rs}} \frac{\partial B_{rs}}{\partial b_{ij}} = \left(\frac{\partial x_i}{\partial X_r} \frac{\partial x_j}{\partial X_s} - \frac{1}{3} g_{rs} \delta_{ij} \right) \frac{\partial F}{\partial B_{rs}}.$$

Thus, in equation (2.28), we can without loss of generality take f_{rs} to be given by (3.12).

(Received September 22, 1957)

Erratum pour le mémoire de PHAM MAU QUAN, *Inductions électromagnétiques en relativité générale et principe de Fermat*, Vol. 1, 54—80 (1957).

Au 2ème théorème du I.6, p. 65, lire:

Dans un milieu isotrope de constantes diélectrique et magnétique ϵ, μ variables, les rayons électromagnétiques sont des géodésiques de longueur nulle de l'espace riemannien \bar{V}_4 muni de la métrique

$$d\bar{s}^2 = \bar{g}_{\alpha\beta} dx^\alpha dx^\beta \equiv \left(g_{\alpha\beta} - \left(1 - \frac{1}{\epsilon\mu} \right) u_\alpha u_\beta \right) dx^\alpha dx^\beta$$

où $g_{\alpha\beta}$ est le tenseur métrique fondamental et u_α le vecteur vitesse unitaire d'univers définis en chaque point du milieu considéré.

(Manuscrit reçu le 7. octobre 1957)

EDITORIAL BOARD

S. BERGMAN

Applied Mathematics
and Statistics Laboratory
Stanford University, California

R. BERKER

Technical University
Istanbul

L. CESARI

Purdue University
Lafayette, Indiana

L. COLLATZ

Institut für Angewandte Mathematik
Universität Hamburg

J. L. ERICKSEN

The Johns Hopkins University
Baltimore, Maryland

A. ERDÉLYI

California Institute of Technology
Pasadena, California

IRMGARD FLÜGGE-LOTZ

Stanford University
California

HILDA GEIRINGER

Harvard University
Cambridge, Massachusetts

H. GÖRTLER

Institut für Angewandte Mathematik
Universität Freiburg i. Br.

D. GRAFFI

Istituto Matematico „Salvatore Pincherle“
Università di Bologna

A. E. GREEN

King's College
Newcastle-upon-Tyne

J. HADAMARD

Institut de France
Paris

H. LEWY

University of California
Berkeley, California

A. LICHNEROWICZ

Collège de France
Paris

C. C. LIN

Mathematics Research Center-U. S. Army
University of Wisconsin
Madison, Wisconsin

C. LOEWNER

Applied Mathematics
and Statistics Laboratory
Stanford University, California

W. MAGNUS

Institute of Mathematical Sciences
New York University
New York City

G. C. McVITTIE

University of Illinois Observatory
Urbana, Illinois

J. MEIXNER

Institut für Theoretische Physik
Technische Hochschule Aachen

C. MIRANDA

Istituto di Matematica
Università di Napoli

C. MÜLLER

Mathematisches Institut
Technische Hochschule Aachen

A. OSTROWSKI

Mathematische Anstalt der Universität
Basel

R. S. RIVLIN

Division of Applied Mathematics
Brown University
Providence, Rhode Island

M. M. SCHIFFER

Stanford University
California

E. STERNBERG

Division of Applied Mathematics
Brown University
Providence, Rhode Island

W. TOLLMIEH

Max-Planck-Institut
für Strömungsforschung
Göttingen

A. TONOLO

Seminario Matematico
Università di Padova

C. TRUESDELL

Mathematics Research Center-U. S. Army
University of Wisconsin
Madison, Wisconsin

H. VILLAT

47, bd. A. Blanqui
Paris XIII

CONTENTS

LIN, C. C., Note on a Class of Exact Solutions in Magneto-hydro-dynamics	391
BERNSTEIN, B., & J. L. ERICKSEN, Work Functions in Hypo-elasticity	396
CRIMINALE, WILLIAM O., JR., J. L. ERICKSEN & G. L. FILBEY, JR., Steady Shear Flow of Non-Newtonian Fluids	410
RABENSTEIN, ALBERT L., Asymptotic Solutions of $u^{iv} + \lambda^2(zu'' + \alpha u' + \beta u) = 0$ for Large $ \lambda $	418
SCHRÖDER, JOHANN, Störungsrechnung bei Eigenwert- und Verzweigungsaufgaben	436
PIPKIN, A., & R. S. RIVLIN, Note on a paper "Further Remarks on the Stress-Deformation Relations for Isotropic Materials"	469
Errata	
Correction to A. E. GREEN & R. S. RIVLIN, The Mechanics of Non-Linear Materials with Memory, Vol. 1, 1—21 (1957)	470
Erratum pour le mémoire de PHAM MAU QUAN, Inductions électromagnétiques en relativité générale et principe de Fermat, Vol. 1, 54—80 (1957)	470